0

2023ko abendua

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

INPAKTU SOZIALERAKO IKERLANEN AKTA

JOURNAL OF RESEARCH FOR SOCIAL IMPACT

INPAKTA

INPAKTA

INPAKTU SOZIALERAKO IKERLANEN AKTA
JOURNAL OF RESEARCH FOR SOCIAL IMPACT

Euskal Herriko Unibertsitatean argitaratua. Leioa (Bizkaia) - Argitalpen Zerbitzua
Edited in Leioa, Bizkaia by the University of the Basque Country (UPV/EHU) - Editorial Service


eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea


Mondragon Unibertsitatea


Universidad de Navarra


IVERSITÉ
DE PAU ET DES
PAYS DE L'ADOUR


upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa


iker
EUSKARA ETA EUSKAL TESTUEN
IKERKETA ZENTROA
CENTRE DE RECHERCHE
SUR LA LANGUE ET
LES TEXTES BASQUES


Orkestra
LEHIAKORTASUNERAKO
EUSKAL INSTITUTUA
DEUSTO FUNDAZIOA


K
AGIRRE LEHENDAKARIA CENTER
for Social and Political Studies


Arantzazulab
GIZARTE BERRIKUNTZA
SOCIAL INNOVATION


TELESFORO
MONZON
eLab


Deusto
Universidad de Deusto
Deustuko Unibertsitatea


ueu


BRTA
BASQUE RESEARCH
& TECHNOLOGY
ALLIANCE


ADItech
COORDINADOR SINAI
Sistema Navarro de I+D+i


elhuyar
ezagutuz aldatzea


Université
BORDEAUX
MONTAIGNE


BOISE STATE UNIVERSITY
BASQUE STUDIES


Center for
Basque Studies


université
de BORDEAUX


CENARRUSA
FOUNDATION for BASQUE CULTURE


SOZIOLINGUISTIKA
KLUSTERRA
Euskara biziberritzeko
ikergunea


EUSKO
IKASKUNTZA
Asmoz ta Jakitez

Dibulgazio-laburpenen erredakzioa:
Elhuyar
Zelai Haundi, 3
Osinalde industrialdea
20170 Usurbil (Gipuzkoa)
zientzia@elhuyar.eus
www.elhuyar.eus

Marka Irudiaren diseinua:
GRAFIK
Proiektu grafikoak. Proyectos gráficos. Graphic projects
grafik.eus

Euskal Herriko Unibertsitateko argitalpen zerbitzua
UPV/EHU Press

Argitalpen zerbitzua
Biblioteka eraikina, lehenengo solairua
Euskal Herriko unibertsitatea
48940 Leioa

Editorial Office
Library Building, 1th floor.
University of the Basque Country
E-48940 Leioa - Spain

emaila: info.editorial@ehu.eus
www.ehu.eus/argitalpenak

ISSN: 2990-2479
Lege gordailua/Depósito legal: LG BI 00790-2023
Printed in Spain

INPAKTA
INPAKTU SOZIALERAKO IKERLANEN AKTA
JOURNAL OF RESEARCH FOR SOCIAL IMPACT

Jatorrizko artikuluak/Original articles

INPAKTA

Iñaki Antiguedad editorearen argazkia

Ur-baliabideen kudeaketa: iturburuetako landarediak baldintzatzen du akuiferoen birkargatze-maila

Klima-larrialdiaren testuinguruan, datozen hamarkadetan lehentasunezkoa izango da lurpeko akuiferoen ur-erreserbak ondo kudeatzea. Azken ikerketek erakutsi dute akuiferoen inguruan dauden iturburuetako landarediak akuiferoen birkarga baldintzatu dezakeela: abeltzaintza-larreek eta baso helduek, ur gutxi kontsumitzen dutenez, akuiferoen birkarga errazten dute; baso-landaketa berriek, ordea, ur-kontsumo handia dutenez, eragotzi. Baso-kudeaketan paradigma-aldaketa ekar dezake honek.

Klima-aldaketaren proiekzioek aurreikusten dute iberiar penintsulan asko murriztuko dela ibaietako ur-emaria udan, lehorte gogorrek eta goi-mendietako elurra goizegi urtzeak eraginda. % 10-30 bitartean murriztuko dira ur-baliabideak mendearen erdialderako. Ondorioz, estrategikoa izango da lurpeko akuiferoak ahal den guztia birkargatzeko neurriak hartzea.

Ikertzaileek ikusi dute sistema hidrologikoak eta lurzorua estaltzen duen landaredia zuzenean lotuta daudela: deforestazioak eta baso-berritzeek inpaktu zuzena dute lurpeko akuiferoen birkarga-prozesuan. Larreak eta 100 urtetik gorako baso helduak dituzten iturburuek errazago birkargatzen dituzte akuiferoak, zuhaitz gazteez estalita daudenen aldean.

«Baso-masa handitze hutsak akuiferoen birkarga zaildu dezake, eta abeltzaintza larreek, aldiz, lagundu»

Izan ere, zuhaitzek, ebapotranspirazio-prozesuan, lurzoruko ura jasotzen dute sustraietatik eta urlurrun moduan askatzen dute, hostoetatik. Prozesu horrek lurzoruko hezetasuna kentzen du, eta akuiferoak birkargatzeko gaitasunean eragiten du. Zuhaitzek larreek baino ur gehiago kentzen diote lurzoruari, batez ere gazteaz direnean eta hazkuntza betean daudenean. Baso lehorrek ere ur asko kontsumitzen dute. Lainoa harrapatzen duten baso hezeek, ordea, akuiferoen birkarga eraginkorra lortzen dutela ikusi da. Lanaren egileek ikusi dute lurzoruaren estaldura aldatze hutsarekin erregimen hidrikoa sasoikoa izatetik iraunkorra izatera pasatu daitekeela, edo alderantziz.

Duela gutxi arte, herrialde askok jarraitutako baso-kudeaketak baso-masa handitzea hobesten zuen kasu guztietan, lurzoruaren higaduratik eta uholdeetatik babesteko estrategia eraginkorrena delako. Hidrologikoki sentikorrak diren arroetan, ordea, lehentasunezkoa izan daiteke akuiferoak birkargatzea erraztuko duen landaredia hobestea.

Mendiko artzaintza desagertzearen ondorioak

Mendi-eremu hezeak ganaduarentzako larre gisa erabili izan dira tradizioz, baina, gaur egun, abeltzaintza galtzen ari da eremu askotan. Ikerketek erakusten dute artzaintza-larreak galdutako eremuetan akuiferoen birkarga murriztu egin dela. Eragin bera du goi-mendietan zuhaitzak basoaren muga naturaletik gora landatzeak eta izurriteak ugaritzeak. Gune horietan guztietan artzaintzarako larreak berreskuratzeak maila freatikoa handitzea ekartzen duela ikusi da nazioartean.

Honek guztiak paradigma-aldaketa bat ekar dezake basoen eta baliabide hidrikoen kudeaketan. Ikertzaileek uste dute tokian tokiko lehentasunezko zerbitzu ekosistemikoei begira hartu behar direla erabakiak: eremu batzuetan, lurzorua higaduraren aurka babesteak izango du lehentasuna, eta, beste batzuetan, ur-akuiferoak betetzeak. Hidrologikoki sentikorrak diren arroetan, larreak berreskuratzea eta baso zaharrak babestea izango da lehentasuna.

InpAkta 00 (2023) 6-17
https://doi.org/10.1387/InpAkta.24999

Vulnerability of water resources to changes in land cover and climate

Francesc Gallart*
Instituto de Diagnóstico Ambiental y Estudios del Agua (IDAEA), CSIC

ABSTRACT: In various hydrographic basins in Iberia, reductions in annual flows attributed to the increase in forest cover have been observed in the last 50-60 years, as much or more important than the changes attributed to variations in precipitation and temperature. Several studies in diverse climates show that groundwater recharge is generally lower under forest cover than under grass or agricultural cover.
The predictions of the climate models coincide in more marked increases in temperature and decreases in precipitation in summer than in winter in most part of Iberia, inducing not only a decrease of water resources but also an increase in water stress of the vegetation and a reduction in dry season flows.
It will be necessary to implement land cover strategies adaptive to climate change scenarios to care for water resources, such as reversing the reforestation of mountain pastures and improving the management of timberlands to reduce their water stress and consumption.

1. Recent evolution of water resources in Iberian basins

During the last decades, various works have been published that show clear decreases in water contributions in the headwaters of several hydrographic basins during the last 50 years, such as the Ebro (García-Ruiz *et al.*, 2001; Gallart & Llorens 2003 and 2004, Beguería *et al.*, 2003, López-Moreno *et al.*, 2011; Buendía *et al.*, 2016a and 2016b), the northern slope of the Duero (Morán-Tejeda *et al.*, 2012; Pisabarro *et al.*, 2019), the south-western tributaries of the Duero (Ceballos-Barbancho *et al.*, 2008), the Llobregat and the Ter (Gallart *et al.*, 2011). In all these studies except in Ceballos-Barbancho *et al.* (2008), it is evident that the decrease in contributions is not justified by climatic reasons, but must be attributed totally or partially to the increase in evapotranspiration caused by the extension of forest cover in these headwaters, since they have not

* Corresponding author / Harremanetan jartzeko: Francesc Gallart. Instituto de Diagnóstico Ambiental y Estudios del Agua (IDAEA), CSIC. Jordi Girona 18 (08034 Barcelona). e-mail: francesc.gallart@idaea.csic.es. Contributed by Iñaki Antigüedad editoreak egindako ekarpena.
How to cite / Nola aipatu: Gallart, Francesc. «Vulnerability of water resources to changes in land cover and climate», *InpAkta* 00 (2023) 6-17 (https://doi.org/10.1387/InpAkta.24999)

experienced any increase of water consumption for irrigation.

Lopez-Moreno *et al.* (2010) also showed that during this period there is a decrease in the frequency and size of floods in the Pyrenean headwaters that is not justified by changes in precipitation but rather must be attributed to the increase in forest cover. They also found an increase in low flows, especially in winter and spring; this increase should be attributed mainly to a premature melting of the snow due to an increase in temperature (0.3º C per decade according to López-Moreno *et al.*, 2011).

It is important to note that both in the headwaters of the Ebro (Gallart & Llorens 2004) and Llobregat and Ter (Gallart *et al.*, 2011), the relationship found between the increase in forest cover and the decrease in contributions is very close to the that can be predicted as a result of the increase in forest cover according to the results obtained in experimental basins throughout the world, synthesized in the equation of Zhang *et al.* (2001). Table 1 summarizes the changes observed in these basins. The gross variation of the water contributions refers to the variation observed in the gauging and the net variation refers to the residual variation of the contributions once the effect of climatic variability has been taken into account, as well as the increase in consumptive uses (irrigation) in the case of the Ebro basin.

These results show that the predictions on the availability of water resources for the future cannot be estimated only from the results of the climate models, but must also take into account the changes suffered by the vegetation cover, which, in turn, they depend on the evolution of the climate, the vegetative cycles and the management of the territory. By way of comparison, the estimated reductions over a period of 50 years due to the increase in forest cover with respect to the average annual flows are from 10 to 17% for the Ebro (Gallart & Llorens 2003), 25% for the Cardener (tributary of the Llobregat) and 12% for the upper Ter (Gallart *et al.*, 2011). Currently there are adequate models available for forecasting water inputs at the operational basin scale (resource management) in the face of changes in cover and climate (Delgado *et al.*, 2010; Zhang *et al.*, 2001; Buendía *et al.*, 2016b; Yang *et al.*, 2021; Karki *et al.*, 2023).

Table 1
Changes in water supply volumes and forest cover in some watersheds. Net variation refers to the residual when climate variability has been cancelled. The periods of study are diverse. Based on Gallart & Llorens (2003) and Gallart *et al.* (2011)

Basin	Water volumes			Forest cover Annual variation (% of basin area)
	Annual mean hm^3	Gross annual variation (%)	Net annual variation (%)	
Ebro	12900	−0.63	−0.20	0.22
Cardener	99	−0.82	−0.50	0.61
Alto Llobregat	210	−1.10	−0.66	0.39
Alto Ter	493	−0.25	−0.26	0.36

2. Effects of afforestation and deforestation on groundwater recharge

In addition to the observations mentioned in operational catchments, experimental results obtained in various environments show that soils tend to be wetter in grassy clearings than under tree cover (Gallart *et al.*, 1997, Gray *et al.*, 2002, Cubera & Moreno 2007, García-Estríngana *et al.*, 2013), which implies higher aquifer recharge rates in the absence of forest cover (Joffre & Rambal 1993, Hatton & George 2000, Ladekarl *et al.*, 2005, Green *et al.*, 2006, Kim & Jackson, 2012).

A marked rise of the water table is a general rule in several large territories elsewhere, where shrub-forest covers have been cleared for grazing or dry-farming (Walker *et al.*, 1999; Scanlon *et al.*, 2005; Leblanc *et al.*, 2008). This is only in some particular cases, such as the localized recharge of karstic systems in semi-arid environments, where change in forest cover may be of little importance (Bazan *et al.*, 2013).

Numerous studies carried out in experimental basins show that low or low flows increase by the same or greater percentage than flood flows when forest cover is eliminated, leading in some cases to a change in the water regime, which goes from seasonal to permanent, or vice versa in the case of afforestation (Hornbeck *et al.*, 1993, Scott *et al.*, 2000, Sikka *et al.*, 2003, Silberstein *et al.*, 2003, Brown *et al.*, 2005). In some experimental or larger basins, the effects of cover changes on low flows are negligible, which has been attributed to a more important effect of geological characteristics than of cover on low flows: if the aquifers are of small volume, they can be completely filled during the wet season whatever the cover, so that low-water flows are not affected by their changes (Robinson *et al.*, 2003, Calder 2005).

3. Expectable changes in hydrological systems over the coming decades

Climate change projections have a high degree of uncertainty, due to the various emission scenarios and the characteristics of the different models. In addition, precipitation, which is the main input to hydrological systems, is not yet reliably simulated by current climate models. Despite these uncertainties, there is a broad consensus among the various models that the increase in temperature and, secondly, the decrease in precipitation will cause a significant reduction, between 10 and 30% of water resources in the Iberian Peninsula for the middle of this century compared to the period of the late 20th century (Arnell *et al.*, 2003, Milly *et al.*, 2005). Later assessments of the IPCC do not significantly modify these projections but show medium confidence in the increase in recently observed and projected droughts (IPCC 2022).

From a seasonal point of view, the predictions show a remarkable agreement that the increases in temperature and decreases in precipitation will be much more pronounced in summer than in winter by the end of this century (Christensen *et al.*, 2007). To simplify the problem, the two main limiting factors of evapotranspiration can be considered: energy (where there is no shortage of water) and water (where there is no shortage of energy); in environments where evapotranspiration is limited by energy, such as high mountains, an increase in temperature in summer can represent a direct increase in evapotranspiration and a decrease in water inputs. Conversely, in environments where evapotranspiration is limited by the availability of water, an increase in temperature in summer may not have an appreciable effect on actual evapotranspiration and water flows. Therefore, it can be expected that the increase in aridity will have greater effects on the water stress of terrestrial plant communities than on the generation of water resources.

However, it can be foreseen that both the increase in the severity of the summer drought and the premature melting of the snow cover in high mountains due to the increase in temperatures, will cause a decrease in dry season flows (García-Ruiz *et al.*, 2011). At the same time, it is foreseeable that there will be a notable increase in the frequency and magnitude of extreme droughts, such that droughts that had a recurrence period of 100 years at the end of the 20th century may have recurrence periods between 70 and less than 10 years in the middle of this century (Lehner *et al.*, 2005). The modifications of floods are less clear in the predictions for the Mediterranean area; although it is generally considered that extreme events will play an increasing role as a result of enhanced atmospheric circulation, an increase in flooding related to climate warming has not been observed (Kundzewicz *et al.*, 2007; IPCC2022).

On the other hand, changes in climate forcing will not occur on a stationary vegetation cover, but rather it will undergo modifications that will affect the hydrological consequences of changes in precipitation and temperature. Some of the causes of these modifications of the vegetal cover are the expansion and growth of forest masses on areas of crops

or abandoned pastures (Poyatos *et al.*, 2003), the afforestation of high mountain areas close to the altitudinal limit of the forest favored by the climate warming (Peñuelas *et al.*, 2007), the decline of species maladapted to changing climatic conditions due to direct causes (Keenan *et al.*, 2011) or due to proliferation of pests (Rouault *et al.*, 2006), and the forest management actions that will be carried out.

4. Water resources and adaptive forest management in the present context

Until the end of the 20th century, forest management carried out in many countries, including Spain, was based on the hypothesis that increasing the mass and forest cover in a hydrographic basin was the best option to protect soils against erosion and regulate the hydrological response. That is to say, to reduce floods and improve the regularity of resources, by favoring the recharge of aquifers and increasing low flows (MAGRAMA 2002). At present, there is sufficient evidence to affirm that forest cover can provide good protection against erosion and small and moderate floods, but at the expense of greater evapotranspiration, which usually leads to appreciable decreases in total water inputs, the aquifer recharge and low water flows. Only forests that capture significant volumes of fog or primeval forests (more than 100 years old) seem to deviate from this general rule (see next section). This paradigm shift has already been collected in documents in the field of forest management, such as the FAO (2006) and the European Forest Institute (2011).

Due to the changing climatic conditions and the paradigm shift, the design of hydrologically sensible forest restoration actions in hydrographic basins and the very definition of forest hydrological ecosystem services are currently at a crossroads and pose serious scientific and technical challenges. The protection against soil erosion and the protection of water resources through afforestation are no longer equivalent, but in most cases they are antagonistic. In certain areas, soil protection will be the preferred criteria for actions, while in others it will be runoff generation or aquifer recharge. The selection of tree species with diverse water needs (Zabaleta *et al.*, 2018), the design of actions such as cutting cycles, thinning or control of the undergrowth, as well as forecasts on the progression, growth and viability of forest masses in different locations of the basins under changing climatic conditions are some of these challenges.

The forest management of each perimeter will depend on the priority ecosystem services, which may be the generation of water resources in any of them. By way of example, the installation of dense forest covers should be avoided in areas where the main water resources are generated in the basins. The headwaters in high and medium humid mountain areas have traditionally been used as pastures for livestock, but currently they tend to experience the expansion of the forest due to the abandonment of livestock activities and global warming, which is producing a decrease of the generation of water resources. The recharge areas of large aquifers are also areas that should be kept free of dense forest cover, provided that this is compatible with other priority ecosystem services.

Forest thinning (density reduction) is a current forest management option that decreases competition for water and light between trees and may increase aquifer recharge and stream flows. In water stressed areas, thinning increases the water availability for the remaining trees but may not decrease the areal water consumption (Gracia *et al.*, 1999). In wetter environments, thinning can increase water yield and groundwater recharge but must be repeated every 3 to 9 years in order to maintain these effects (Del Campo *et al.*, 2022).

Finally, as is often the case in various semi-arid areas, the excessive spread of phreatophyte stands should be avoided to conserve groundwater resources and low water flows (Doody *et al.*, 2011).

5. Complementary information

5.1. Role of vegetation cover changes in water balances

If only the limiting factors to evapotranspiration are considered, the differences between tree and short vegetation covers can be simplified (Calder 1998b).

When the limiting factor is energy, a tree canopy typically has a lower albedo (captures more radiative energy) and much higher aerodynamic roughness (exchanges energy and matter with a much thicker atmospheric layer) than an herbaceous canopy. During rainy events, trees have a greater aerial biomass, so they can intercept a greater volume of water and, for the above reasons, they have a greater capacity to evaporate the intercepted water. When the limiting factor is water, trees often have significantly deeper root systems than herbaceous plants, so they have access to a greater volume of water. As a result, trees have a greater capacity to evapotranspire water than herbaceous covers under a wide range of conditions, which explains the coincidence of results obtained in various parts of the world (see for example Brown *et al.*, 2005, Zhang *et al.*, 2011).

If a forest cover implies a greater real evapotranspiration than a herbaceous cover, it will cause a higher atmospheric humidity and therefore a greater return of this humidity in the form of precipitation (see an explanation of the mechanisms in Millán *et al.*, 2005). What usually happens in reality is that this greater precipitation becomes effective outside the scope of the same basin, on a very large or continental basin scale. There are several observed and simulated examples of continental-scale precipitation variation as a result of changes in vegetation cover (Calder 1998a, 2005; Sant'Anna Comar *et al.*, 2023).

There are few exceptions to the general pattern that forest cover causes decreased watershed inputs compared to herbaceous cover. In some areas with recurring mist, trees can capture part of this mist and convert it into "hidden precipitation", which can represent significant water input in areas with little conventional rainfall (Bruijnzeel *et al.*, 2005). It has also been observed that when primeval forests are felled or burned, runoff shows an initial increase, but drops sharply with respect to the original value as young forest develops. It has therefore been suggested that the water consumption of some forests gradually decreases with their age, until reaching stable rates not much higher than those of a herbaceous cover, when they reach an age of 100-150 years (Kuczera 1987, Vertessy *et al.*, 2001).

5.2. *Current perspectives on the management of water resources in river basins*

The traditional management of water resources in a basin is based on estimating the contributions of water available in rivers and aquifers, using gauging data or simulations carried out with a hydrological model, and assigning flows to the various intended uses (MARM 2008). The big problem with this approach from the point of view of sustainability, especially in a context of global change, is that it assumes that the hydrological cycle begins in (and can only be managed from) the river or the reservoir. The evidence that the uses and vegetation covers of the headwaters of the basins have a relevant role in the water contributions of the rivers has shown the need to carry out an integral management of the water balance of the basins, taking into account all the consumptions (natural and artificial) in particular when water resources in the basin are limited.

The Republic of South Africa has played a pioneering role in the development of this comprehensive approach when it became apparent that the establishment of commercial forestry in the headwaters of the watersheds diminished the water inputs necessary for subsistence crops in the lower parts. This led to the enactment of the "National Water Act" (Republic of South Africa, 1998), in which commercial forest plantation is considered "Stream Flow Reduction Activity" and is required to financially compensate to the authority of the basin for the withdrawal of flows that they cause. In the words of Calder (2005), this law represented a departure from the colonial rules and regulations of humid European countries, which favoured the interests of a dominant group with privileged access to land and water.

Other market-based instruments have been elsewhere introduced, based on the payment of off-site beneficiaries (water users) for the water provision services provided through the Payment for Ecosystem Services scheme (PES),(European Forest Institute (2011). This may allow the establishment of water -producing land uses that are not economically sustainable, such as extensive livestock farming.

To facilitate integrated land and water management, in 1993 the concept of "green water" was introduced at an FAO seminar (1995) to express and quantify the water consumed by terrestrial ecosystems, including meadows and dryland crops , reserving the term "blue water" for liquid runoff or groundwater. It is estimated that, globally, two thirds of precipitation water is used in the production of biomass of terrestrial ecosystems (green water) while only one third is blue water; 60-70% of the world's food production is produced thanks to "green water" that is not taken into account in traditional water balances.

This approach allows the allocation of water resources to all uses of the basin, whether irrigated or not, making it possible to integrate water management with land management (Calder 2005, Falkenmark & Rockstorm 2004). Changes in use and vegetation cover, even if they are not irrigated, can thus be managed by the basin authority. This facilitates, for example, the establishment of economic compensation for ecosystem services related to changes in use or forest management actions, such as felling or thinning (reducing tree density).

Figure 1 shows a plot of the rainfall partition between green water and blue water for a range of precipitation; Figure 1B shows that in dry and sub-humid climates the change in vegetation cover represents a large relative change in water partitioning.

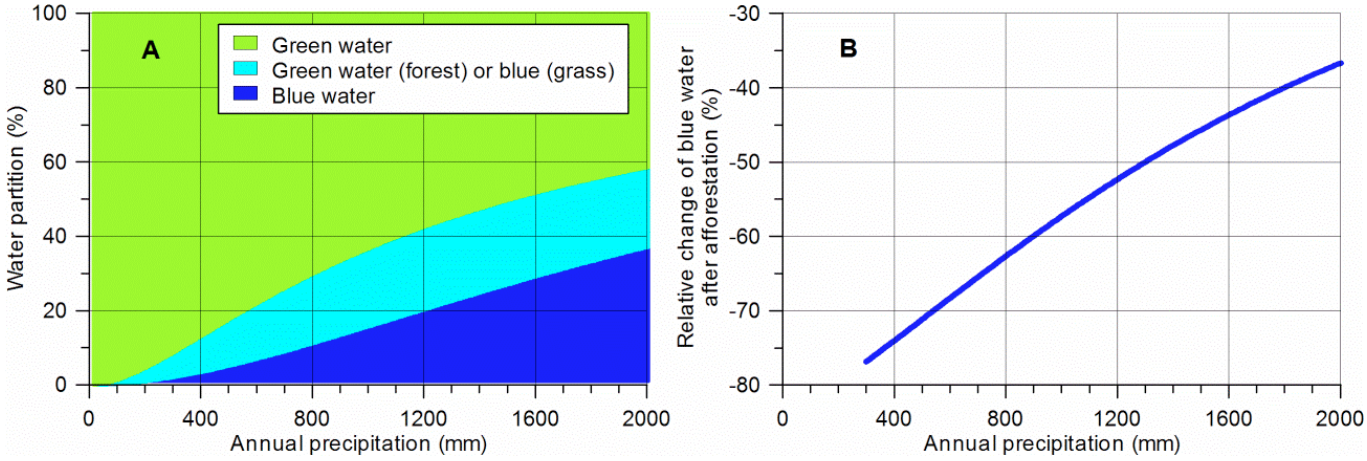


Figure 1: Diagram of the partition of rainwater in a range of precipitation between: i) green water, ii) green water if the vegetation cover is forest or blue water if it is herbaceous and iii) blue water in any case. B: Simulation of the relative change in runoff in the case of total afforestation of a basin with herbaceous cover. The original equation of Zhang *et al.* (2001) has been used, which assumes a subtropical climate with high evapotranspiration demand (1100 mm per year)

5.3. Hydrological models

The series of water flows observed at gauging stations are usually not suitable for hydrological planning. This is due to two causes: the observed flows are affected by the management of the resources (reservoirs, intakes and returns), and they are only available at the observation points and not at the points of interest for management. For these reasons, hydrological (or rainfall-runoff) models are used to generate flow series "in natural regime" from climatic records at points of interest in the basin.

The models used for this purpose in Spain are normally the Sacramento (Burnash 1995), which is calibrated using the gauging series of stations that are considered to be little influenced by uses, and the SIMPA (Ruiz-García 1999) which is parameterized with maps of terrain features. In both cases, the models are used to convert the precipitation and potential evapotranspiration series into input series, assuming that the characteristics of the basins do not vary over time. For this reason, the contributions simulated in this way should be properly called flows "in climatic regime" instead of "natural regime".

The comparison of the simulated contributions in climatic regime with the observed contributions is one of the methods that allow studying the hydrological role of the changes suffered in the headwaters, once discounting the changes in consumptive uses of water in the event that they have occurred. The use of the records and simulations carried out by the authorities responsible for management has the double advantage of being more convincing to these bodies and of facilitating the application of the results for management.

There are several hydrological models designed to study the hydrological role of vegetation cover changes. The simplest model is the equation of Zhang *et al.* (2001) that expresses the annual water balance in an area of uniform vegetation cover using a single empirical parameter (type of vegetation cover). Although for the development of this equation only basins in subtropical climates free of snowfall were used, it has given acceptable results when applied to basins with Pyrenean headwaters with a minor role of snow in total precipitation (Delgado *et al.*, 2010 , Gallart *et al.*, 2011). This equation can be easily implemented in a Geographic Information System to study the response of medium and large basins to various scenarios (Bradford *et al.*, 2001).

The daily scale model most used in practice is SWAT (Arnold & Fohrer 2005) which, although it has remarkably weak conceptual bases, is based on a structure of "hydrological response units" and has the great advantage of having an extensive community of users who exchange information and experiences. As default parameterisations of SWAT usually produce unrealistic predictions of the hydrological responses of forest covers, advanced parameterisations have been recently developed (Haas *et al.*, 2022; Karki *et al.*, 2023).

The HYLUC model (Calder *et al.*, 2003, Delgado *et al.*, 2010) is a daily scale sub-basin aggregate model specially designed for this problem, which has strong conceptual foundations. The distributed TETIS model has also been used at daily scale to analyse both the hydrological consequences of land cover change and climatic variations by comparing a sequence of different land-use maps with a constant land use scenario (Buendía *et al.*, 2016b).

Lastly, there are also other models, such as the monthly SIMPA model, which, although not designed for this purpose, admit the adaptation of its parameters to different vegetation covers to simulate changes (Ruiz-García 1999).

When using any type of model for this purpose, it is advisable to verify that the change in vegetation cover parameters essentially represents a change in the real evapotranspiration rates, and that the other changes in the simulated flows and stores are consistent with this change.

6. References

Arnell NW (2003) Effects of IPCC SRES emissions scenarios on river runoff: a global perspective. *Hydrology and Earth System Sciences* 7: 619-641.

Arnold JG, Fohrer, N (2005) SWAT2000: current capabilities and research opportunities in applied watershed modelling. Hydrological Processes 19 (3): 563-572

Bazan RA, Wilcox BP, Munster C, Gary M 2013. Removing woody vegetation has little effect on conduit flow recharge. Ecohydrology, 6 (3): 435-443.

Beguería S, López-Moreno JI, Lorente A, Seeger M, García-Ruiz JM (2003) Assessing the effects of climate oscillations and land-use changes on streamflow in the Central Spanish Pyrenees. Ambio 32: 283-286.

Bellot J, Benet A, Sanchez JR, Chirino E (2001) Likely effects of land use changes on the runoff and aquifer recharge in a semiarid landscape using a hydrological model. Landscape and Urban Planning 55 (1): 41-53

Bellot J, Chirino E (2013) Hydrobal: An eco-hydrological modelling approach for assessing water balances in different vegetation types in semi-arid areas. Ecological Modelling 266: 30-41.

Bradford A, Zhang L, Hairsine P (2001) Implementation of a mean annual water balance model within a GIS framework and application to the Murray-Darling basin. Cooperative Research Centre for catchment Hydrology. Technical Report 01/8. Bruce, Australia. Available: http://www.nfp-facility.org/5865-06c5b4f-34ca7eb9155af198cbfdf42498.pdf. Accessed 27 January 2014.

Brown AE, Zhang L, McMahon TA, Western T, Vertessy RA (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. Journal of Hydrology 310: 28-61.

Bruijnzeel LA, Eugster W, Burkard R (2005) Fog as a hydrologic input. En: Anderson M, McDonnell J, Anderson M, editores. Encyclope-

dia of Hydrological Sciences. John Wiley and Sons, Chichester, pp. 559-582

Buendia C, Batalla R J, Sabater S, Palau A, & Marcé R. (2016a). Runoff trends driven by climate and afforestation in a Pyrenean Basin. Land degradation & development, 27(3), 823-838.

Buendia C, Bussi G, Tuset J, Vericat D, Sabater S, Palau A, & Batalla R J (2016b). Effects of afforestation on runoff and sediment load in an upland Mediterranean catchment. Science of the Total Environment, 540, 144-157.

Burnash RJC (1995) The NWS River Forecast System - Catchment Modeling. En: Singh VP, editor. Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, Colo. pp: 311-366.

Calder IR (1998a) Water-resource and land-use issues. SWIM Paper 3. Colombo, Sri Lanka: International Water Management Institute. Available: http://www.iwmi.cgiar.org/Publications/SWIM_Papers/PDFs/SWIM03.PDF. Accessed 27 January 2014

Calder IR (1998b) Water use by forests, limits and controls. Tree Physiology 18(8-9): 625-631.

Calder IR (2003) Assessing the water use of short vegetation and forests: development of the Hydrological Land Use Change (HYLUC) model. Water Resources Research 39 (11): 1318.

Calder IR (2005) Blue Revolution: Integrated Land and Water Resources Management. Earthscan, London, 353pp.

Ceballos-Barbancho A, Morán-Tejeda E, Luengo-Ugidos M Á, & Llorente-Pinto M. (2008). Water resources and environmental change in a Mediterranean environment: the south-west sector of the Duero river basin (Spain). Journal of Hydrology, 351(1-2), 126-138.

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional Climate Projections. En: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (editores.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Clarke CJ, George RJ, Bell RW, Hatton TJ (2002) Dryland salinity in South-Western Australia: its origins, remedies and future research directions. Australian Journal of Soil Research 40: 93-113.

Cubera E, Moreno G (2007) Effect of land-use on soil water dynamic in dehesas of Central–Western Spain. Catena 71: 298-308.

Delgado J, Llorens P, Nord G, Calder IR, Gallart F (2010) Modelling the hydrological response of a Mediterranean medium-sized headwater basin subject to land cover change: The Cardener River basin (NE Spain). Journal of Hydrology, 383 (1-2): 125-134.

Doody TM, Nagler PL, Glenn EP, Moore GW, Morino K, Hultine KR, Benyon RG (2011) Potential for water salvage by removal of non-native woody vegetation from dryland river systems. Hydrological Processes 25 (26): 4117-4131.

European Forest Institute (2011) Agua para los bosques y la Sociedad en el Mediterráneo. Un difícil equilibrio. Birot, I, Gracia, C. Palahí M. (Editores). What Science can tell us. European Forest Institute. Available: http://www.efi.int/files/attachments/publications/efi_what_science_can_tell_us_1_2011_sp.pdf. Accessed 27 January 2014.

Falkenmark M, Rockstorm J (2004) Balancing Water for Humans and Nature. The New Approach in Ecohydrology. Earthscan, Sterling, VA. USA. Available: http://f3.tiera.ru/1/genesis/655-659/658000/a342d7a0a312513181fec0537faaadbe. Accessed 27 January 2014.

FAO (1995) Land and water integration and river basin management. Rome, Italy. Available: http://www.fao.org/docrep/v5400e/v5400e00.HTM Accessed 27 January 2014.

FAO 2006. The new generation of watershed management programmes and projects. Food And Agriculture Organization of the United Nations, Rome. Available: ftp://ftp.fao.org/docrep/fao/009/a0644e/a0644e.pdf. Accessed 27 January 2014.

Gallart F, Llorens P (2003) Catchment management under environmental change: Impact of land cover change on water resources. Water International 28 (3): 334-340

Gallart F, Llorens P (2004) Observations on land cover changes and water resources in the headwaters of the Ebro catchment, Iberian Peninsula. Physics and Chemistry of the Earth 29 (11-12): 769-773

Gallart F, Latron J, Llorens P, Rabadà D (1997) Hydrological functioning of Mediterranean mountain basins in Vallcebre, Catalonia: Some challenges for hydrological modelling. Hydrological Processes 11: 1263-1272

Gallart F, Delgado J, Beatson SJV, Posner H, Llorens P, Marcé R (2011) Analysing the effect of global change on the historical trends of water resources in the headwaters of the Llobregat and Ter river basins (Catalonia, Spain). Physics and Chemistry of the Earth, Parts A/B/C 16 (13): 655-661

Garcia-Estríngana P, Latron J, Llorens P, Gallart F (2013) Spatial and temporal dynamics of soil moisture in a Mediterranean mountain area (Vallcebre, NE Spain). Ecohydrology 6 (5): 741-753

Garcıa-Ruiz J ., Begueria S, López-Moreno J I, Lorente-Grima A & Seeger M. (2001). Los recursos hıdricos superficiales del Pirineo aragonés y su evolución reciente: Geoforma Ediciones, Logroño, 191, ISBN: 848777931X

García-Ruiz JM, López-Moreno JI, Vicente-Serrano SM, Lasanta-Martínez T, Beguería S (2011) Mediterranean water resources in a global change scenario. Earth Sciences Reviews 105: 121-139.

Gracia CA, Sabaté S, Martínez JM, Albeza E. (1999). Functional Responses to Thinning. In: Rodà, F., Retana, J., Gracia, C.A., Bellot, J. (eds) Ecology of Mediterranean Evergreen Oak Forests. Ecological Studies, vol 137. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-58618-7_23

Gray AN, Spies TA, Easter MJ (2002) Microclimatic and soil moisture responses to gap formation in coastal Douglas-fir forests. Canadian Journal of Forest Research 32: 332-43

Green JC, Reid I, Calder IR, Nisbet TR (2006) Four-year comparison of water contents beneath a grass ley and a deciduous oak wood overlying Triassic sandstone in lowland England. Journal of Hydrology 329(1-2): 16-25

Hatton T & George R. (2000). The Role of Afforestation in Managing Dryland Salinity. In: Plantations, Farm Forestry and Water. Proceedings

of a national workshop, 20-21 July 2000, Melbourne. RIRDC Publication No. 01/20. ISSN 1440-6845. Kingston, ACT, 28-35.

Hornbeck JW, Adams MB, Corbett ES, Verry ES, Lynch JA (1993) Long-term impacts of forest treatments on water yield: a summary for northeastern USA. Journal of Hydrology 150 (2/4): 323-344

IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp., doi:10.1017/9781009325844.

Joffre R, Rambal S (1993) How Tree Cover Influences the Water Balance of Mediterranean Rangelands. Ecology 74 (2): 570-582

Keenan T, Serra JM, Lloret F, Ninyerola M, Sabate S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO_2 matters! Global Change Biology 17: 565-579

Kim JH, & Jackson RB. (2012). A global analysis of groundwater recharge for vegetation, climate, and soils. Vadose Zone Journal, 11(1).

Kuczera G (1987) Prediction of water yield reductions following a bushfire in ash mixed species eucalypt forest. Journal of Hydrology 94: 215-236

Kundzewicz ZW, Mata LJ, Arnell NW, Döll P, Kabat P, Jiménez B, Miller KA, Oki T, Sen,Z, Shiklomanov IA (2007) Freshwater resources and their management. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds., Cambridge University Press, Cambridge, UK, 173-210.

Ladekarl UL, Rasmussen KR, Christensen S, Jensen KH, Hansen B (2005) Groundwater recharge and evapotranspiration for two natural ecosystems covered with oak and heather. Journal of Hydrology, 300 (1-4): 76-99

Leblanc MJ, Favreau G, Massuel S, Tweed S, Loireau M, Cappelaere B (2008) Land clearance and hydrological change in the Sahel: SW Niger. Global and Planetary Change, 61(3-4): 135-150

Lehner B, Döll P, Alcamo, J, Henrichs H, Kaspar F (2005) Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated assessment. Climatic Change, 75: 273-299

López-Moreno JI, Beguería S, García-Ruiz JM (2010) Trends in high flows in the central Spanish Pyrenees: response to climatic factors or to land-use change?. Hydrological Sciences Journal, 51(6): 1039-1050

López-Moreno JI, Vicente-Serrano SM, Morán-Tejeda E, Zabalza J, Lorenzo-Lacruz J, García-Ruíz JM (2011) Impact of climate evolution and land-use changes on water yield in the Ebro basin. Hydrology and Earth System Sciences 15: 311-322

McCulloch JSG, Robinson M (1993) History of forest hydrology. Journal of Hydrology 150: 189-216

MAGRAMA, 2002. Plan Forestal Español. Available: http://www.magrama.gob.es/es/biodiversidad/temas/politica-forestal/pfe_tcm7-30496.pdf. Accessed: 27 January 2014.

MARM (2008) Orden ARM/2656/2008 de 10 de septiembre por la que se aprueba la instrucción de planificación hidrológica. Boletín Oficial del Estado 229:38472-38582.

Milly PCD, Dunne KA, Vecchia AV (2005) Global pattern of trends in streamflow and water availability in a changing climate. Nature 438, 347-350

Morán-Tejeda E, Ceballos-Barbancho A, Llorente-Pinto JM, López-Moreno JI (2012) Land-cover changes and recent hydrological evolution in the Duero Basin (Spain). Regional Environmental Change 12:17-33

Peñuelas J, Ogaya R, Boada M, Jump AS (2007) Migration, invasion and decline: changes in recruitment and forest structure in a warming-linked shift of European beech forest in Catalonia (NE Spain). Ecography 30:829-837

Pisabarro A., Pellitero R, Serrano E & Lopez-Moreno JI (2019). Impacts of land abandonment and climate variability on runoff generation and sediment transport in the Pisuerga headwaters (Cantabrian Mountains, Spain). Geografiska Annaler: Series A, Physical Geography, 101(3), 211-224.

Poyatos R, Latron J, Llorens P, (2003) Land use and land cover change after agricultural abandonment. The case of a mediterranean mountain area (Catalan Pre-Pyrenees). Mountain Research and Development 23(4): 362-368

Republic of South Africa (1998) National Water Act. (Act No. 36 of 1998). Government Gazette 398 no. 19182.

Robinson M, Cognard-Plancq AL, Cosandey C, David J, Durand P, Fuhrer HW, Hall R, Hendriques MO, Marc V, McCarthy R, McDonnell M, Martin C, Nisbet T, O'Dea P, Rodgers M, Zollner A (2003) Studies of the impact of forests on peak flows and baseflows: a European perspective. Forest Ecology and Management, 186: 85-97.

Rouault G, Candau JN, Lieutier F, Nageleisen LM, Martin JC, Warzée N (2006) Effects of drought and heat on forest insect populations in relation to the 2003 drought in Western Europe. Annals of Forest Sciences 63: 613-24.

Ruiz García JM (1999) Modelo distribuido para la evaluación de recursos hídricos. Monografías del CEDEX, 67. Ministerio de Fomento, Madrid.

Sant'Anna Commar LF, Medeiros Abrahao G & Heil Costa M. (2023). A possible deforestation-induced synoptic-scale circulation that delays the rainy season onset in Amazonia. Environmental Research Letters 18 044041

Scanlon BR, Reedy RC, Stonestrom DA, Prudic DE, Dennehy KF (2005) Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. Global Change Biology, 11(10): 1577-1593

Scott DF, Prinsloo FW, Moses G, Mehlomakulu M, Simmers ADA (2000) A re-analysis of the South African afforestation experimental data, Water Research Commission, WRC Report No. 810/1/00, South Africa.

Segura-Beltrán F, Sanchis-Ibor C (2013) Assessment of channel changes in a Mediterranean ephemeral stream since the early

twentieth century. The Rambla de Cervera, eastern Spain. Geomorphology 201: 199-214

Sikka AK, Samra JS, Sharda VN, Samraj P, Lakshmanan V (2003) Low flow and high flow responses to converting natural grassland into bluegum (Eucalyptus globulus) in Nilgiris watersheds of South India. Journal of Hydrology 270: 12-26

Silberstein R, Adhitya A, Dabrowski C (2003) Changes in flood flows, saturated area and salinity associated with forest clearing for Agriculture. Technical Report 03/1. CRC Centre for Catchment Hydrology, Monash, Australia. Available: http://www.catchment.crc.org.au/pdfs/technical200301.pdf. Last accessed 27 October 2008

Vertessy RA, Watson RGR, O'Sullivan SK (2001) Factors determining relations between stand age and catchment water balance in mountain ash forests. Forest Ecology and Management 143: 13-26

Walker GR, Gilfedder M & Williams J. (1999). Effectiveness of current farming systems in the control of dryland salinity. CSIRO Land and Water. ISBN 0 643 06556 3

Yang Z, Hou F, Cheng J & Zhang Y. (2021). Modeling the Effect of Different Forest Types on Water Balance in the Three Gorges Reservoir Area in China, with CoupModel. Water, 13(5), 654.https://doi.org/10.3390/w13050654

Zabaleta A, Garmendia E, Mariel P, Tamayo I, & Antigüedad I. (2018). Land cover effects on hydrologic services under a precipitation gradient. Hydrology and Earth System Sciences, 22(10), 5227-5241.

Zhang L, Dawes WR, Walker GR (2001) Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water Resources Research, 37: 701-708.

Cristina Peña editorearen argazkia

Ibaietatik itsasora iristen den zaborra kudeatzeko gakoak

Ibaietatik itsasora iristen den zaborra kudeatzea oso prozesu konplexua da. Parte bat itsaso zabalera iristen da, eta beste parte bat itsasertzean pilatuta geratzen da. Zaborren dinamikan eragiten duten faktore nagusiak identifikatu dituzte: ibaiaren inguruko lurren erabilera, ozeanoaren korronteak, fenomeno meteorologikoak, zaborraren flotagarritasuna eta urtaroa. Kutsadura handiko ibaietan faktore hauek guztiak monitorizatuta edukitzea lagungarri izango da agintarientzat larrialdiko erantzunak emateko estrategiak diseinatzeko.

Ibaietako zaborra kudeatzea erronka garrantzitsua da itsasoko eta kostaldeko ekosistemak babesteko. Itsasorako bide naturalak izanik, zaborraren kudeaketa eskasagatik ibaietara iristen diren arrasto guztiek itsasoan barreiatuta amaitzen dute. Ikerketek erakutsi dute ibaiek ekarritakoa dela ozeanoan dagoen plastikoaren % 80, eta hiriguneetako ibai txikiak seinalatzen dituzte erantzule gisa. Europak berak 307-925 milioi zabor flotagarri askatzen ditu ozeanoan urtero, hiriguneetako ibaien bidez.

Europako itsasoetan, Bizkaiko Golkoaren hego-ekialdea da itsasoko zabor flotagarria pilatzen den gune esanguratsuenetako bat. Ez da erraza arazoa kudeatzea, ez baitago argi zaborra nondik eta nola datorren. Udalerriek kudeaketa-irizpide argiak izan ditzaten, zortzi ibaietako zaborraren dinamika aztertu da ikerketa honetan: Deba, Urola, Oria, Urumea, Oiartzun, Bidasoa, Urdazuri eta Aturri. Zaborraren ibilbidea monitorizatzeko, frekuentzia altuko lurreko radarrak eta buia flotagarriak erabili dira.

Zaborraren dinamika konplexua

Konplexutasun handia du zaborraren dinamika ulertzeak, faktore askok baldintzatzen baitute. Hasteko, ibaien formak, sakonerak eta ur-emariak. Baina inguruko lurren erabilerak eta biztanle-dentsitateak ere eragingo du ibairatzen duten zaborraren izaera eta kantitatean. Era berean, Euskal Herriaren mendebaldeko ibaiek 7-24 km-ko plataforma kontinental estua izaten dute itsasoratzen direnean, baina 70 km-ko plataforma zabala Landetara hurbildu ahala. Plataforman sortzen diren korronteak ere intentsitate eta iraupen handikoak izaten dira neguan eta udazkenean, eta txikikoak udaran.

Konplexutasun horren barruan, ondorio argi batzuk atera dituzte ikertzaileek: gure ibaiek garraiatzen duten zabor flotagarriaren % 95 plastikoa da. Haien artean, ohikoagoak dira flotagarritasun txikikoak (esaterako, plastikozko poltsak) flotagarritasun handikoak baino (hala nola botilak). Haizeak batez ere flotagarritasun handiko zaborra mugiarazten du; beraz, euskal kostaldean itsasoratzen den zaborrean eragin handiagoa dute azaleko korronte ozeanikoek haizeak berak baino.

Flotagarritasun handiko zaborrak oso azkar mugitzen dira haizearen bidez, baina udaran partikulen % 97 itsasertzean harrapatuta geratzen dira astebeteren buruan. Udazkenean, % 54ra murrizten da trabatuta geratzen den zaborren tasa. Flotagarritasun txikiko zaborren kasuan, ordea, % 25 baino gutxiago geratzen dira harrapatuta. Gehienak itsaso zabalean barreiatzen dira.

Ondorioz, zabor-kontzentrazio altuenak Euskal Herriaren ekialdeko muturrean eta udaran hautematen dira, eta flotagarritasun handiko zaborrez osatua egoten da.

Ikertzaileek argi dute ezen, ibaietako eta itsasoko zaborrak desagerrarazteko, ezinbestekoa dela lurreko zaborra murriztu eta tokiko hondakinen kudeaketa hobetzea, politika eta araudi zorrotzagoak ezarriz. Horrekin batera, ibaietako zaborra itsasora ez iristeko neurriak hartzeko gomendatzen dute. Alabaina, jada itsasoratzen ari den zaborra modu eraginkorrean jasotzeko, ezinbestekoa ikusten dute ikerketa honetan identifikatutako dinamika kontuan hartzea.

InpAkta 00 (2023) 18-41
https://doi.org/10.5194/os-18-1703-2022

Modelling floating riverine litter in the south-eastern Bay of Biscay: a regional distribution from a seasonal perspective

Irene Ruiz[1], Anna Rubio[1], Ana J. Abascal[2] and Oihane C. Basurko[1]

[1] AZTI, Marine Research, Basque Research and Technology Alliance (BRTA); [2] IHCantabria - Instituto de Hidráulica Ambiental de la Universidad de Cantabria

ABSTRACT: Although rivers contribute to the flux of litter to the marine environment, estimates of riverine litter amounts and detailed studies on floating riverine litter behaviour once it has reached the sea are still scarce. This paper provides an analysis of the seasonal behaviour of floating marine litter released by rivers within the south-eastern Bay of Biscay based on riverine litter characterizations, drifters, and high-frequency radar observations and Lagrangian simulations. Virtual particles were released in the coastal area as a proxy of the floating fraction of riverine litter entering from rivers and reaching the open waters. Particles were parameterized with a wind drag coefficient (C_d) to represent their In contrast, low-buoyancy items took longer to arrive at the shoreline, particularly during spring with fewer than 25% of particles beached by the end of the simulations. The highest concentrations (> 200 particles km^{-1}) were recorded during summer for C_d 4% in the French region of Pyrénées-Atlantiques. Results showed that the regions in the study area were highly affected by rivers within or nearby the region itself. These results couple observations and a river-by-river modelling approach and can assist decision-makers on setting emergency responses to high fluxes of floating riverine litter and on defining future monitoring strategies for heavily polluted regions within the south-eastern Bay of Biscay.

1. Introduction

Rivers act as key vectors bringing improperly disposed and mismanaged litter from land into the marine environment. Riverine litter poses a large threat to freshwater systems by degrading aquatic life, impacting freshwater quality, and increasing economic losses linked with human activities (van Emmerik and Schwarz, 2020; Al-Zawaidah *et al.*, 2021). However, most of the litter research conducted to date has focused on marine environments (87%) when compared to freshwaters systems (13%)

* **Corresponding author / Harremanetan jartzeko:** Irene Ruiz. AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Pasaia, 20110, Spain. e-mail: iruiz@azti.es. **Contributed by** Cristina Peña **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Ruiz, Irene *et al.* «Modelling floating riverine litter in the south-eastern Bay of Biscay: a regional distribution from a seasonal perspective», *Ocean Science* | (2022)18:1703-1724 (https://doi.org/10.5194/os-18-1703-2022)

(Blettler *et al.*, 2018). Indeed, riverine litter contributions to oceans are still uncertain, and results vary depending on the input data and the model applied (Lebreton *et al.*, 2017; Schmidt *et al.*, 2017; Mai *et al.*, 2020). Recent findings derived from extensive modelling efforts suggest that about 1600 rivers worldwide account for 80% of plastic inputs to the ocean, with small urban rivers among the most polluting (Meijer *et al.*, 2021). Models require comprehensive field data and consistent and harmonized protocols to validate the amounts, type, and size of riverine inputs (González-Fernández and Hanke, 2017; Wendt-Potthoff *et al.*, 2020; Margenat *et al.*, 2021). Such comprehensive data were obtained in Europe thanks to the RIMMEL project (González-Fernández and Hanke, 2017). This research concluded that between 307 and 925 million floating riverine litter items are annually transferred into the ocean mainly through small rivers, streams, and coastal runoff (González-Fernández *et al.*, 2021).

Once it has reached the sea, floating riverine litter can accumulate close to the shoreline or it can be transported to open waters, reaching even remote areas far from the coast. Indeed, the distribution and fate of floating litter in the marine environment is affected by the metocean conditions (currents, turbulence, wind) but also by the buoyancy of the objects (Ryan, 2015; Lebreton *et al.*, 2019; Maclean *et al.*, 2021). Objects with low buoyancy are mainly driven by currents, by contrast with highly buoyant items, which are driven along the water surface partially by winds. This wind effect ("windage") on floating marine litter behaviour has been further investigated by Lagrangian modelling studies in the open ocean (Allshouse *et al.*, 2017; Maximenko *et al.*, 2018; Lebreton *et al.*, 2019; Abascal *et al.*, 2009) when compared to the coastal area (Critchell and Lambrechts, 2016; van Utenhove, 2019; Tong *et al.*, 2021). The lack of observational data is one of the key limitations for parameterizing the windage effect and accurately predict floating marine litter behaviour. However, observations derived from drifting buoys, such as those provided for decades by the Global Drifter programme, have been used to fill this gap. They have allowed simulating more realistic floating marine litter pathways from origin to fate by integrating experimental windage parameterizations and the corresponding comparison between observed and modelled trajectories (Duhec *et al.*, 2015; Pereiro *et al.*, 2018; Rizal *et al.*, 2021). Nowadays, more affordable and environmentally friendly solutions are gaining force among researchers, as drifters are built using biopolymers (Novelli *et al.*, 2017; D'Asaro *et al.*, 2020) or have compact and lightweight designs with a GPS-tracking component for easy deployment (Meyerjürgens *et al.*, 2019; van Sebille *et al.*, 2021).

At the coastal scale, windage parameterization combined with realistic knowledge on coastal circulation becomes crucial to reduce the uncertainties of modelled trajectories (Van Sebille *et al.*, 2020). Land-based high-frequency radar systems (hereafter HF radars, Rubio *et al.*, 2017) offer the opportunity to monitor surface currents in coastal areas, where the transport processes are significantly more complex than in open-ocean waters due to the effect of the coast, the bathymetry, and other local forcings (e.g. river discharges or coastal upwellings). In the south-eastern Bay of Biscay (hereafter SE Bay of Biscay), as part of the operational oceanography system EuskOOS, an HF radar provides near-real-time surface currents fields. The system has already been used to study surface coastal transport processes in the area in combination with multisource data (Manso-Narvarte *et al.*, 2018, 2021; Rubio *et al.*, 2011, 2013, 2018, 2020; Solabarrieta *et al.*, 2014, 2015, 2016). The HF radar is also a good example of effective monitoring of surface currents with strong potential for floating marine litter management. Research conducted by Declerck *et al.*, (2019) in the SE Bay of Biscay provided the first assessment of floating marine litter transport and distribution in the region, coupling surface current observations from the EuskOOS system, Lagrangian modelling, and riverine inputs. Nowadays, these observations are used by local authorities both in real time and in hindcast in the framework of the operational service FML-TRACK to collect floating marine litter in the area. However, the accurate modelling of the transport and fate of floating marine litter needs to consider the vari-

ety of floating objects and sources and additional physical parameterization, such as windage.

This paper aims at estimating the seasonal behaviour of the floating marine litter fraction released by rivers within the SE Bay of Biscay reaching open waters. To do so, a Lagrangian model was forced by real observations from the EuskOOS HF radar and particles were parameterized to represent floating marine litter trajectories of two groups of items according to their buoyancy. Riverine litter collected from a local barrier was characterized at the laboratory to explore the fraction of high- and low-buoyancy items. Since most of the items were low-buoyancy, simulations considering only surface currents were performed as the reference. Complementary Lagrangian simulations for highly buoyant items (and less abundant in the area) were also performed. In this case, four low-cost buoys with a similar buoyancy to certain highly buoyant items were built and released in three different rivers. Drifter data were used to parameterize the wind effect on this type of item and consequently achieve more accurate results.

2. Study area

The study was conducted in the SE Bay of Biscay, between north-eastern Spain (Basque Country) and south-western France (Landes). The study area extends from 43.27 to 44.58°N and from 3.18 to 1.27° W, falling within the coverage area of the HF radar station of the operational oceanography system EuskOOS (Fig. 1). The study area comprises two Basque regions (Bizkaia (Spain) and Gipuzkoa (Spain)), two French departments (Pyrénées-Atlantiques and Landes), and eight rivers (Deba (Spain), Urola (Spain), Oria (Spain), Urumea (Spain), Oiartzun (Spain), Bidasoa (Spain), Nivelle (France), and Adour (France)). The mean annual river discharge varies widely between rivers from 3.71 $m^3 s^{-1}$ (Oiartzun) to 350 $m^3 s^{-1}$ (Adour) (Sheppard,2018), and the population density differs between the Spanish and French border: 44.8 inhabitants km^{-2} (Landes) to 303.7 inhabitants km^{-2} (Basque Country) (Eurostat, 2019). The bathymetry in the SE Bay of Biscay is characterized by the presence of a narrow continental shelf ranging 7 and 24 km wide in the Basque area, gradually increasing along the French coast up to about 70 km (Bourillet *et al.*, 2006; Rodríguez *et al.*, 2021). The continental shelf in the SE Bay of Biscay comprises two main areas: the Aquitaine shelf with a N-S orientation and the Cantabrian shelf with an E-W orientation. The continental slope is very pronounced, with a slope of the order of 10%-12% (Sheppard, 2018). Over the continental shelf, the ocean circulation is marked by seasonal variability. At shorter temporal scales, circulation in the study area is mostly modulated by the bathymetry and the coastal orientation, the density-driven currents, and winds (Le Boyer *et al.*, 2013; Solabarrieta *et al.*, 2014). Tidal currents are quite weakly constrained by the topography and the width of the continental shelf (Lavin *et al.*, 2006; González *et al.*, 2007; Karagiorgos *et al.*, 2020). Along-shelf currents are more intense and persistent during winter and autumn (about 10-15 cm s^{-1}), contrary to the other seasons, especially in summer (about 2.5 cm s^{-1}) (Charria *et al.*, 2013). In winter, the prevailing SW winds causes an E to N flow from the Spanish coasts towards the French coasts. The moderate to strong NW winds occurring in spring and summer induce a S and SW surface current circulation accompanied by a greater variability (Solabarrieta *et al.*, 2015). In winter, westerly winds in the Basque coast reinforce the slope current (named the Iberian Poleward Current (IPC)), a warm and saline intrusion trapped within 50 km of the shelf edge, reaching its greatest velocities (up to 70 cm s^{-1}) during this season. The IPC favours the along-slope transport of water masses (Solabarrieta *et al.*, 2014; Porter *et al.*, 2016). The exchange between shelf and deep-sea waters in winter is associated with the generation of eddies, from the interaction of currents with the topography (Lavin *et al.*, 2006; Rubio *et al.*, 2018; Teles-Machado *et al.*, 2016). Maximum run-offs combined with SW winds also allow river plumes to spread northwards and along the French shore during winter. However, this path changes in spring, when river discharges are reduced and winds blow from the north-west (Lavin *et al.*, 2006; Puillat *et al.*, 2006).

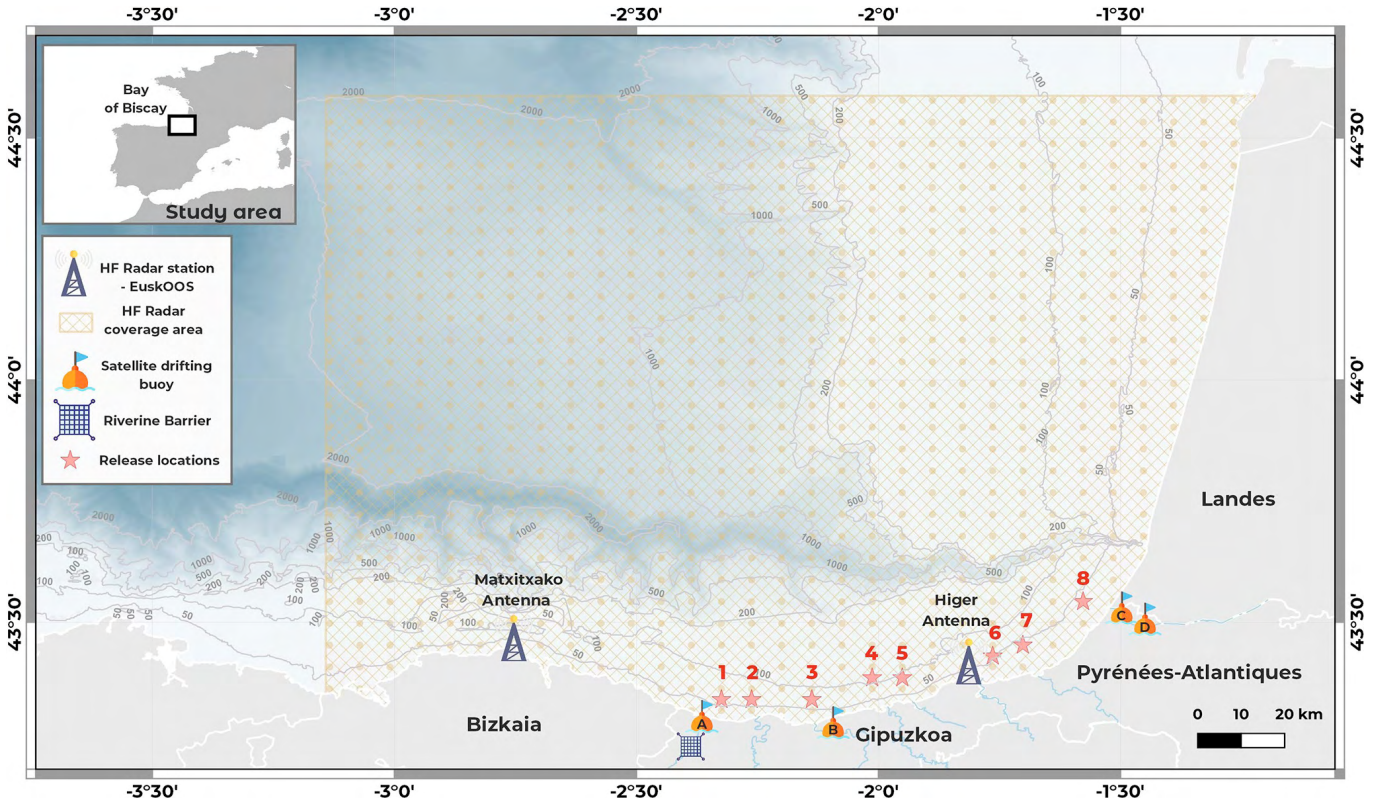


Figure 1: Study area with the release locations of the satellite drifting buoys and the riverine barrier. Dots in light yellow represent the nodes of the HF radar grid. Dots in orange represent the trajectories of the buoys. Numbers with stars in pink correspond to the particle releasing location for floating marine litter simulations: (1) Deba, (2) Urola, (3) Oria, (4) Urumea, (5) Oiartzun, (6) Bidasoa, (7) Nivelle, and (8) Adour rivers

First global modelling studies coupling ocean circulation and Lagrangian particle-tracking models reported that the SE Bay of Biscay is a hotspot for floating marine litter (Lebreton *et al.*, 2012; van Sebille *et al.*, 2012). Recent Lagrangian modelling studies combining measured and predicted surface currents by the HF radar and the Iberian Biscay Irish System (IBI) Copernicus model revealed that floating marine litter circulation in the SE Bay of Biscay is marked by a high seasonal variability. Results showed a higher retention during spring and summer and a northward dispersion along the French coast during autumn and winter (Declerck *et al.*, 2019; Rubio *et al.*, 2020). Surface currents derived from the Regional Ocean Modelling System (ROMS) and a particle-tracking model were combined by Pereiro *et al.* (2019) to track the numerical drifters representing floating marine litter in the Bay of Biscay. In this study, longer residence times and higher concentrations were observed in the SE Bay of Biscay when compared to north-western Iberian coastal waters, particularly in winter. From numerical simulations run using the HYCOM model, Rodríguez-Díaz *et al.* (2020) showed that floating marine litter items with high windage (C_d 3%-5%) tend to accumulate in nearshore areas of the Bay of Biscay or end up beached. This trend is consistent with recent numerical simulations combining surface currents from the operational IBI and the numerical model TESEO that also revealed that the highly buoyant items (C_d 4%) rapidly beach in the SE Bay of Biscay, mainly in spring and summer (Ruiz *et al.*, 2022a). Since June 2020, innovative detection and tracking solutions combining ocean modelling and remote observation systems have been operating in the SE Bay of Biscay to support floating marine litter reduction strategies both downstream (interception at the sea with collection vessels and

on beaches with cleaning facilities) and upstream (source identification and reduction) (Delpey *et al.*, 2021). However, research on floating marine litter behaviour in the SE Bay of Biscay is still in its early stage. Further experiments are needed to fully understand the role of windage, waves, and tides in the complex 3D circulation patterns governing coastal accumulation.

3. Methods and data

3.1. Riverine litter sampling

In spring 2018, a riverine barrier was placed in the Deba River (Gipuzkoa) to retain and collect floating riverine litter during low to moderate flows. This barrier enabled a passive sampling to characterize litter items in the lab. The barrier, which consisted of a nylon artisanal net supported by hard floats (buoys), was 40 m long and 0.6 m high with a 60 mm mesh size (see photos in Appendix A). The sampling was conducted weekly from April to June 2018. In total eight riverine litter samples were collected. Litter items were quantified, weighed, and categorized in the lab according to the master list included in the *Guidance on Monitoring of Marine Litter in European Seas* (Joint Research Centre, 2014). Items were grouped into seven types of material (artificial polymer materials, rubber, cloth/textile, processed/-worked wood, paper/cardboard, metal, and glass/ceramics) and further classified into 44 categories (see the classification in Appendix B). Riverine litter items were also categorized into two groups (low- and high-buoyancy items) considering their exposure to wind based on Ruiz *et al.* (2022a).

3.2. Drifter observations

Four satellite drifting buoys (herein after "low-cost buoys") were built by the authors and deployed one by one in the river mouths of the Deba (Buoy A), Oria (Buoy B), and Adour (buoys C and D) between April 2018 and November 2018 (Fig. 1, Table 1). The low-cost buoys provided positioning every 5 min using satellite technology. Low-cost buoys were 9 cm in height and 9.5 cm in float diameter and weighed approximately 200 g (Fig. 2). A GPS (SPOT Trace device) powered by four AAA cells was placed in the bottom of a high-density polyethylene (HDPE) plastic container sealed to guarantee water tightness. They were chosen because of their capability to ensure a reasonable balance between an accurate signal emission and their purchase and communication fees. SPOT Trace devices have been used over the past few years in coastal and open-ocean applications in a wide range of studies. Studies range from calibrating HF radars (Martínez Fernández *et al.*, 2021) and tracking drifting objects such as icebergs (Carlson *et al.*, 2020), pelagic *Sargassum* (Putman *et al.*, 2020; van Sebille *et al.*, 2021), or fishing vessels (Widyatmoko *et al.*, 2021; Hoenner *et al.*, 2022) to search and rescue training (Russell, 2017) and oil spill and litter monitoring (Novelli *et al.*, 2018; Meyerjürgens *et al.*, 2019). Almost two-thirds of the buoy floated above the water surface, thus preventing any satellite signal losses. Buoys A and D and transmitted their positions on an ongoing basis until their landing. Buoys B and C stopped emitting while they were drifting. In all cases, battery lifetime was enough for an adequate performance of the buoys. Once on land, citizens collected the buoys and reported their corresponding location.

Table 1
Locations, periods, and distances covered by the drifting buoys

Buoy ID	River	Initial date (UTC+1)	Final date	Distance covered (km)
A	Deba	16 September 2018 08:00	4 October 2018 07:00	116.10
B	Oria	12 April 2018 16:00	18 April 2018 12:00	118.72
C	Adour	29 July 2018 20:00	2 August 2018 20:00	71.21
D	Adour	28 November 2018 09:00	30 November 2018 11:00	64.41

Figure 2: Main components of the low-cost buoy. The structure: (a) HDPE container and SPOT Trace device powered by four AAA cells. Assembly process: (b) final appearance once the buoy is sealed; the buoy is labelled with contact information both within and outside; (c) the SPOT Trace was fixed at the base of the container with adhesive tape to avoid twists and turns of the buoy

3.3. *HF radar current observations and wind data*

Surface velocity current fields were obtained from the EuskOOS HF radar station composed by two antennas located at Matxitxako and Cape Higer and covering the SE Bay of Biscay since 2009, a range up to 150 km from the coast. The EuskOOS HF radar is part of JERICO-RI and it is operated following JERICO-S3 project best practices, standards, and recommendations (see Solabarrieta *et al.*, 2016; Rubio *et al.*, 2018, for details). Data consist of hourly current fields with a 5 km spatial resolution obtained from using the gap-filling OMA methodology (Kaplan and Lekien, 2007; Solabarrieta *et al.*, 2021). In total, 85 OMA modes, built setting a minimum spatial scale of 20 km and applied to periods with data from the two antennas, were used to provide maximum spatiotemporal continuity in the HF radar current fields, which is a prerequisite of performing accurate Lagrangian simulations. The application of the OMA methodology has been validated for the Lagrangian assessment of coastal ocean dynamics in the study area by Hernández-Carrasco *et al.* (2018). HF radar velocities were quality controlled using procedures based on velocity and variance thresholds, signal- to-noise ratios, and radial and total coverage, following standard recommendations (Mantovani *et al.*, 2020). Data subsets were built for the Lagrangian simulations avoiding periods with temporal gaps (still present in the case of the failure of one or two antennas) of more than a few hours. Hourly ERA5-U10-wind fields were obtained from the atmospheric reanalysis computed using the IFS model of the European Center for Medium-Range Weather Forecast (ECMWF) (see C3S, 2019 for details). The ERA5 atmospheric database covers the Earth on a 30 km horizontal grid using 137 vertical levels from the surface up to a height of 80 km and provides estimates of a large number of atmospheric, land, and oceanic climate variables, currently from 1979 to within 3 months of real time. Both HF radar current observations and wind data cover the drifter's emission periods and the selected week-long periods between 2009 and 2021 for riverine litter simulations (see Appendix C for the selected periods).

3.4. *Particle transport model*

The application of the transport module of the TESEO particle-tracking model (Abascal *et al.*, 2007, 2017a, b; Chiri *et al.*, 2020) was two-fold: (1) to simulate the transport and fate of floating marine litter entering from rivers and reaching the open waters of the SE Bay of Biscay and (2) to estimate a windage coefficient by calibrating the model according

to the low-cost buoy trajectories. This module allows for simulating passive particles driven by surface currents, wind, and turbulent diffusion. Particle trajectories were calculated using the following equation:

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{u}_a(\boldsymbol{x}_i, t) + \boldsymbol{u}_d(\boldsymbol{x}_i, t), \quad (1)$$

where $\boldsymbol{u}_a$ and $\boldsymbol{u}_d$ are the advective velocity and diffusive velocity, respectively, for the $\boldsymbol{x}_i$ point and t time. The advective velocity is calculated as the lineal combination of the wind and currents according to

$$\boldsymbol{u}_a = \boldsymbol{u}_c + C_d \boldsymbol{u}_w, \quad (2)$$

where $\boldsymbol{u}_c$ is the surface current velocity, $\boldsymbol{u}_w$ is the wind velocity at 10 m over the sea surface, and C_d is the wind drag coefficient. The turbulent diffusive velocity is obtained using Monte Carlo sampling in the range of velocities [$\boldsymbol{u}_d$, $\boldsymbol{u}_d$], which are assumed to be proportional to the diffusion coefficients (Hunter *et al.*, 1993; Maier-Reimer and Sündermann, 1982). For each time step Δt, the velocity fluctuation is defined as

$$|\boldsymbol{u}_d| = \sqrt{\frac{6D}{\Delta t}}, \quad (3)$$

where D is the diffusion coefficient, whose value is 1 $m^2\ s^{-1}$ in accordance with previously modelling work for floating marine litter (Pereiro *et al.*, 2019; Ruiz *et al.*, 2022a). Simulations were forced by HF radar surface current velocity and wind data and interpolated at the particle's position for integrating the trajectories. Beaching along the coast was implemented by a simple approach: if the particle reaches the shoreline, it is identified as beached, and it is removed from the computational process. TESEO has been calibrated and validated by comparing virtual particle trajectories to observed surface drifter trajectories at regional and local scale (Abascal *et al.*, 2009, 2017a, b; Chiri *et al.*, 2019). TESEO is a 3D numerical model conceived to simulate the transport and degradation of hydrocarbons, but it has also been successfully applied to the study of transport and accumulation of marine litter in estuaries (Mazarrasa *et al.*, 2019; Núñez *et al.*, 2019) and in open waters (Ruiz *et al.*, 2022a).

3.4.1. Wind drag estimation

Two simulation strategies were combined for (1) estimating the wind drag coefficient and (2) studying the seasonal behaviour of floating items in the area (Sect. 3.5.2) The wind drag coefficient (C_d) was determined by comparing the observed trajectories provided by the low-cost buoys and the modelled trajectories performed with TESEO. The test was done through different parameterizations of the wind drag coefficient ranging from 0% to 7% (Table 2). This range was chosen based on previously floating marine litter studies coupling Lagrangian modelling and observations from satellite drifting buoys (Carson *et al.*, 2013; Stanev *et al.*, 2019; Van Der Mheen *et al.*, 2019). The coefficient providing the lowest error was considered the best coefficient to simulate highly buoyant litter. Due to the grid limitations of the surface currents and wind data in the coastal area, the comparison was not initialized at the launching position of the low-cost buoys (river mouths), but instead it was initialized at the closest grid element that contained valid currents and wind data (Table 1). Observed positions were interpolated into a uniform 1 h time, fitting the metocean temporal resolution. A release of 1000 virtual particles was performed every 4 h at the corresponding observed position (Table 2). Particles were tracked over a 24 h period and the trajectory of the centre of mass of all the particles was computed at every time step to represent the track of the particle cloud. Observations were compared to modelled trajectories using the simple separation distance, which is the difference between the observed and the computed position of the centre of mass at a time step t. The mean separation distance $\overline{D(t^{\mathrm{mod}})}$ was calculated for every modelled position based on the simple separation distance following Eq. (4):

$$\overline{D(t^{\mathrm{mod}})} = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{X}^{\mathrm{mod}}(t^{\mathrm{mod}}) - \boldsymbol{X}^{\mathrm{obs}}(t^{\mathrm{obs}})\right|, \quad (4)$$

where $\boldsymbol{X}^{\mathrm{mod}}$ t^{mod} and $\boldsymbol{X}^{\mathrm{obs}}(t^{\mathrm{obs}})$ are the modelled and observed trajectories for the simulation period i of a total of N periods. A mean separation distance curve was computed for every wind drag coefficient derived from the mean separation distance curves

of the four buoys. The area beneath the mean separation distance curve was calculated to select the more suitable wind drag coefficient. The area $\tilde{D}$ was calculated as a numerical integration over the forecast period via the trapezoidal method following Eq. (4). This method approximates the integration over an interval by breaking the area down into trapezoids with more easily computable areas:

$$\tilde{D} \approx \int_{t^{\mathrm{mod}=1}}^{t^{\mathrm{mod}=24}} \overline{D\left(t^{\mathrm{mod}}\right)} dt. \quad (5)$$

Table 2
Simulation, release, and physical parameter values for wind drag estimation and floating riverine litter simulations

	Simulation parameters			Release parameters		Physical parameters	
	Number of particles	Integration time	Time step	Release locations	Release time	Turbulent diffusion coefficient	Wind drag coefficient (C_d)
Simulations for wind drag estimation	1000 per location	24 h	60 s	At the observed locations of the buoy	Over the emitting period of the buoy at spaced intervals of 4 h	1 $m^2 s^{-1}$	0%, 2%, 3%, 4%, 5%, 6%, 7%
Seasonal riverine litter simulations	500 per river	1 week	60 s	At a distance of 2.5 nautical miles from the river mouth	At the beginning of the selected time period (10 periods per season)	1 $m^2 s^{-1}$	0%, 4%

3.4.2. Lagrangian seasonal simulation of riverine litter items

Seasonal simulations were run for low- and high-buoyancy items to assess the seasonal differences in the transport and fate of floating riverine litter once it has reached the open waters of the SE Bay of Biscay. Particles were released around 2.5 nautical miles off the shoreline due to the complexity in resolving small-scale processes of floating riverine and marine litter behaviour in and close to the river mouths. As parameterizations concerning wind effect linked to the object characteristics are scarce, the optimal wind drag coefficient estimated for the buoys (see Sect. 3.5.1) was accounted for by simulating the behaviour of the objects highly exposed to wind. No wind drag parameterization (C_d 0%) was applied for low-buoyancy objects not subjected to the wind effect. A total of 10 periods per season uniformly distributed within the study period (2009-2021) were considered for running the simulations based on the availability of HF radar surface current datasets (Appendix C). In total, 80 simulations (40 for C_d 0% and 40 for C_d 4%) were run for 7 d. For each simulation, 4000 particles were released in eight rivers (500 per river) assuming that river discharges are equal despite the seasonal variations and the morphological differences between rivers (Table 2). The total number of particles modelled for C_d 0% was the same as C_d 4%. Post-processing was carried out to compute by river (1) the particles' evolution over the time from their release until their arrival at the shoreline and (2) the particles' distribution on the shoreline, counting the number of beached particles per kilometre of shoreline and indicating the spatial concentration per region.

4. Results

4.1. Riverine litter characterization

In total 1576 items and 11.597 kg of floating riverine litter were sampled and characterized (Fig. 3). Plastic was the most common type of riverine litter in

terms of the number of items (95.1%) and in weight (67.9%); they were also frequent glass/ceramics (16.1%) and cloth/textile items (6.9%) when counted by weight. The top 10 litter items accounted for 93.3% by number and 72.6% by weight of the total riverine litter (Table 3). Plastic/polystyrene pieces between 2.5 and 50 cm and other plastic/polystyrene identifiable items (e.g. food labelling) were the most abundant in terms of number (71.2%) and weight (16.9%). Weakly buoyant items encompassed almost 91% by number and 68% by weight of litter items (Fig. 4).

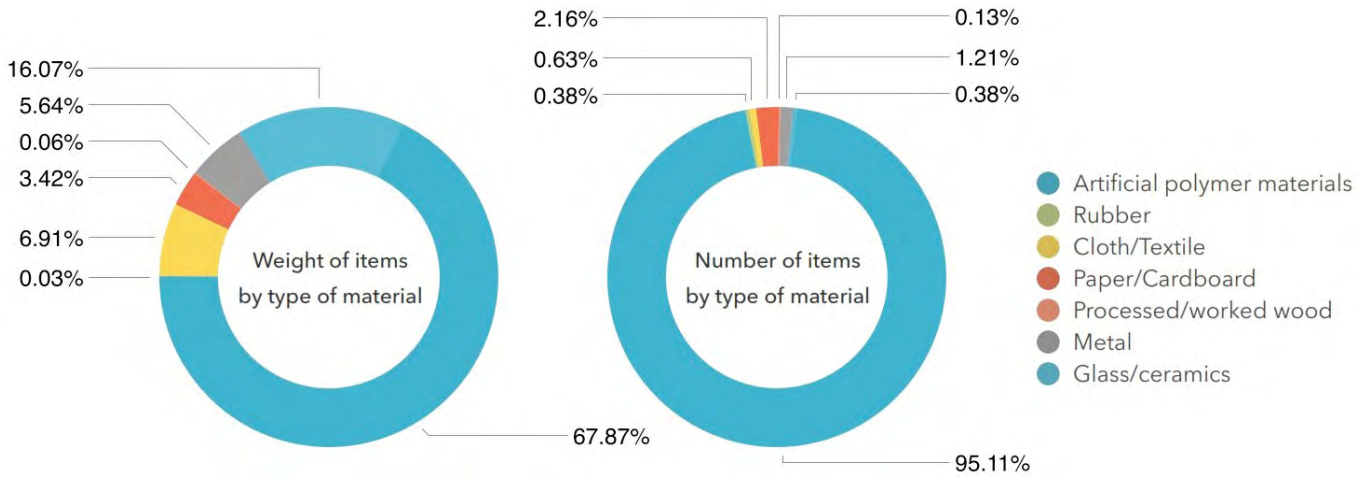


Figure 3: Composition of riverine litter by type of material in terms of the number of items and weight. Items were collected by the barrier placed in the Deba River (Gipuzkoa) between April and June 2018

Table 3

Top 10 (X) riverine litter items collected from the barrier located in the Deba River (Gipuzkoa) between April and June 2018. Items have been ranked by abundance (left) and weight (right) according to the master list categories of beach litter items and classified based on their exposure to the wind effect

Top X by number of items				
Ranking	TSG_ML General code	General name	Number of items (%)	Type of item
1	G76	Plastic/polystyrene pieces 2.5 cm >< 50 cm	71.19%	Low buoyant
2	G10	Food containers incl. Fast food containers	6.21%	Highly buoyant
3	G124	Other plastic/polystyrene items (identifiables)	3.68%	Low buoyant
4	G30	Crips packets/sweet wrappers	3.55%	Low buoyant
5	G20-G24	Plastic caps and lids/Plastic rings	2.41%	Low buoyant
6	G96	Sanitary towels/panty liners/ backing strips	2.22%	Low buoyant
7	G158	Other paper items	1.33%	Low buoyant
8	G5	What remains of rip-off plastic bags	1.33%	Low buoyant
9	G77	Plastic/polystyrene pieces >50 cm	0.82%	Low buoyant
10	G3	Shopping bags incl.pieces	0.51%	Low buoyant
		TOTAL	93.25%	

Top X by weight of items				
Ranking	TSG_ML General code	General name	Weight (%)	Type of item
1	G124	Other plastic/polystyrene items (identifiables)	16.88%	Low buoyant
2	G200	Bottles incl. Pieces	15.80%	Highly buoyant
3	G76	Plastic/polystyrene pieces 2.5 cm >< 50 cm	9.48%	Low buoyant
4	G96	Sanitary towels/panty liners/backing strips	9.48%	Low buoyant
5	G10	Food containers incl. Fast food containers	6.04%	Highly buoyant
6	G135	Clothing (clothes, shoes)	4.16%	Low buoyant
7	G77	Plastic/polystyrene pieces > 50 cm	2.91%	Low buoyant
8	G145	Other textiles (incl.rags)	2.77%	Low buoyant
9	G175-G176	Cans (beverage/food)	2.60%	Highly buoyant
10	G3	Shopping bags incl.pieces	2.52%	Low buoyant
		TOTAL	72.64%	

FLOATING RIVERINE LITTER CLASSIFICATION BY WIND EXPOSURE

	Highly buoyant items	Low buoyant items
Number of items	1,436	140
Weight of items (kg)	7.880	3.717

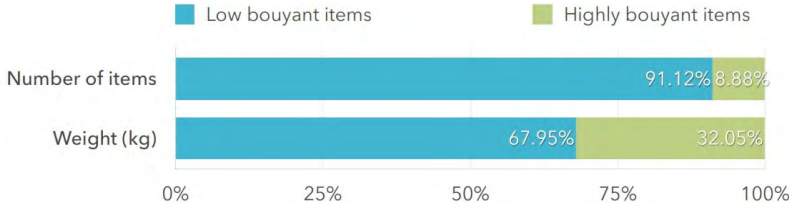


Figure 4: Riverine litter classification based on the exposure to the wind effect. Items were collected from the barrier located in the Deba River (Gipuzkoa) between April and June 2018

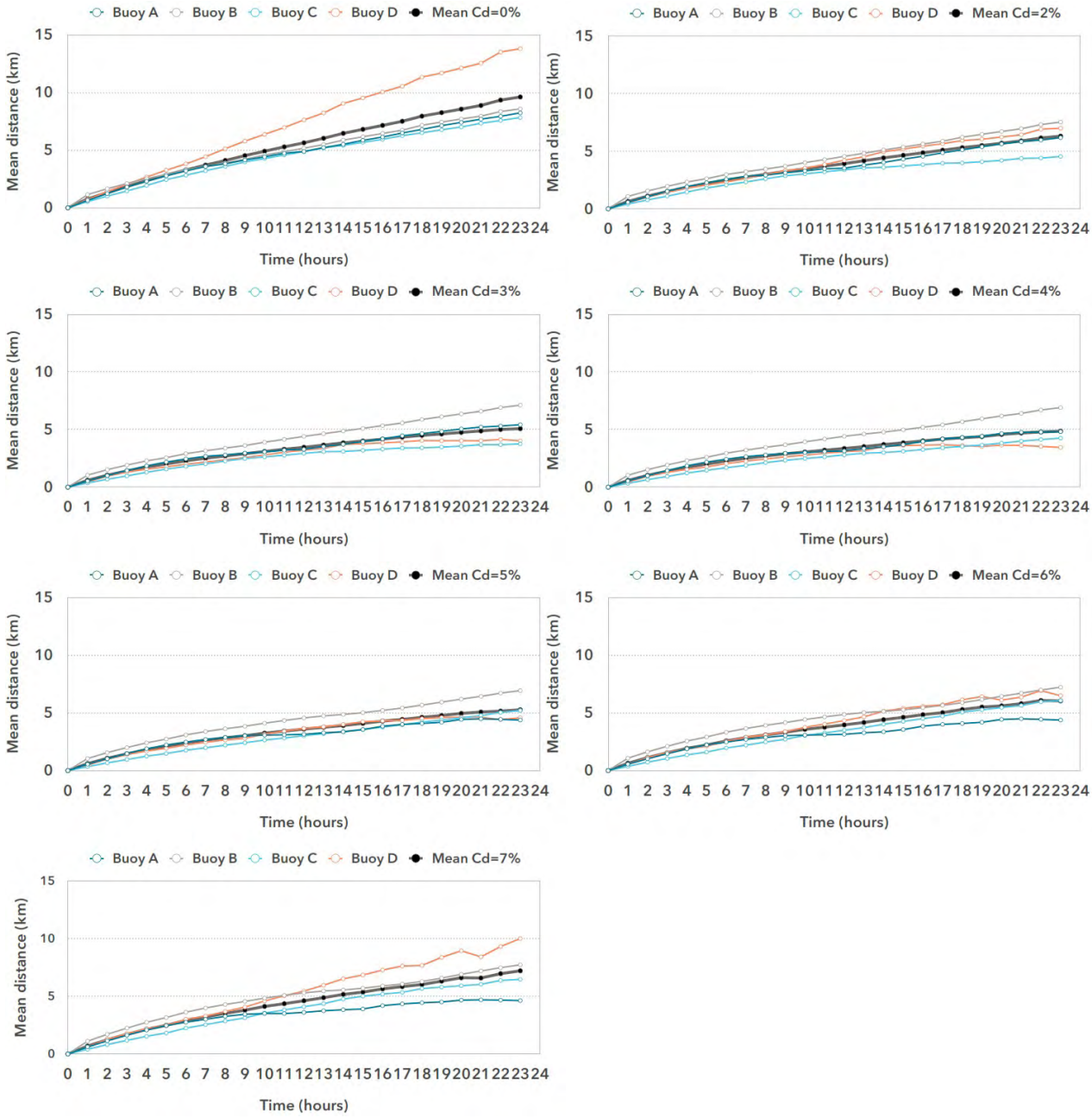


Figure 5: Mean separation distance between modelled and observed trajectories for each wind drag coefficient. The dark line is the mean curve used for the trapezoidal integration

4.2. *Wind drag coefficient for drifting buoys*

Total distances covered by drifting buoys ranged from 62 to 118 km (Table 1), and they all scattered over the HF radar coverage area. Buoys provided their position data over 385 h before beaching on the Landes and Gipuzkoa shore-lines. When compared with numerical trajectories obtained using different C_d parameterizations, the mean separation distance $\overline{D(t^{mod})}$ increased nearly linearly with time for all the parameterizations, achieving a maximum separation of almost 14 km at 24 h for C_d = 0% (Fig. 5). Overall, using no windage parameterization provided the largest $\tilde{D}$. Simulations parameterized with C_d 4% provided the best results with an average standard deviation (SD) of 3.2 1.25 km and a maximum value of 4.85 km at 24 h. When assessing the mean separation distance for all the modelled positions at every observed position of the buoys, the most common range separation distance for C_d 4% was 2-4 km (Fig. 6). Hence, a wind drag coefficient of 4% was applied in the remaining analysis to estimate the behaviour of highly buoyant items.

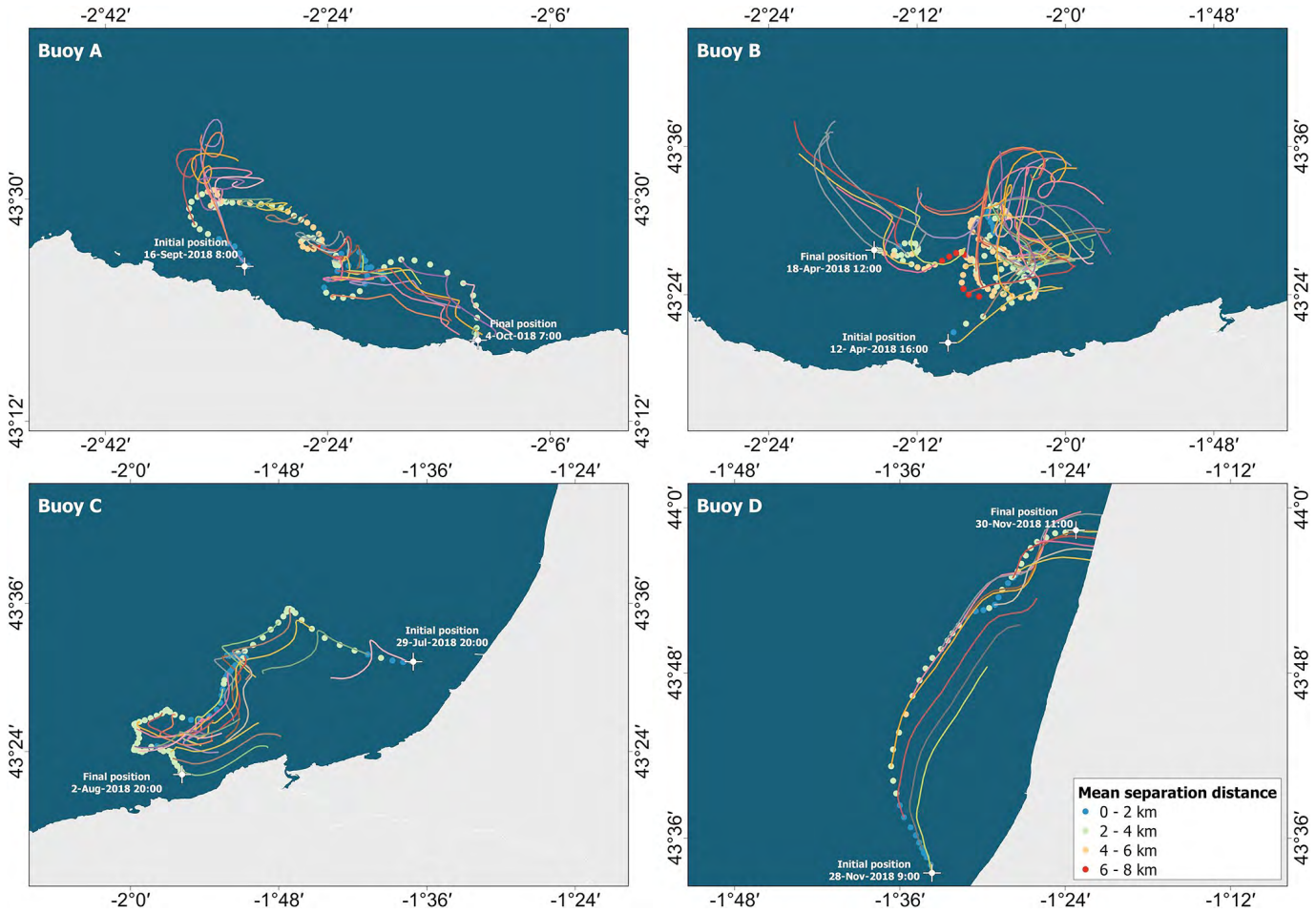


Figure 6: Spatial mean distance between modelled and observed trajectories of buoy A, B, C, and D with a drag coefficient C_d 4%. Particle trajectories were simulated during 24 h, with a re-initialization period every 4 h. The modelled trajectories are shown in solid lines. Circles represent the mean separation distance at the observed position for all the modelled positions

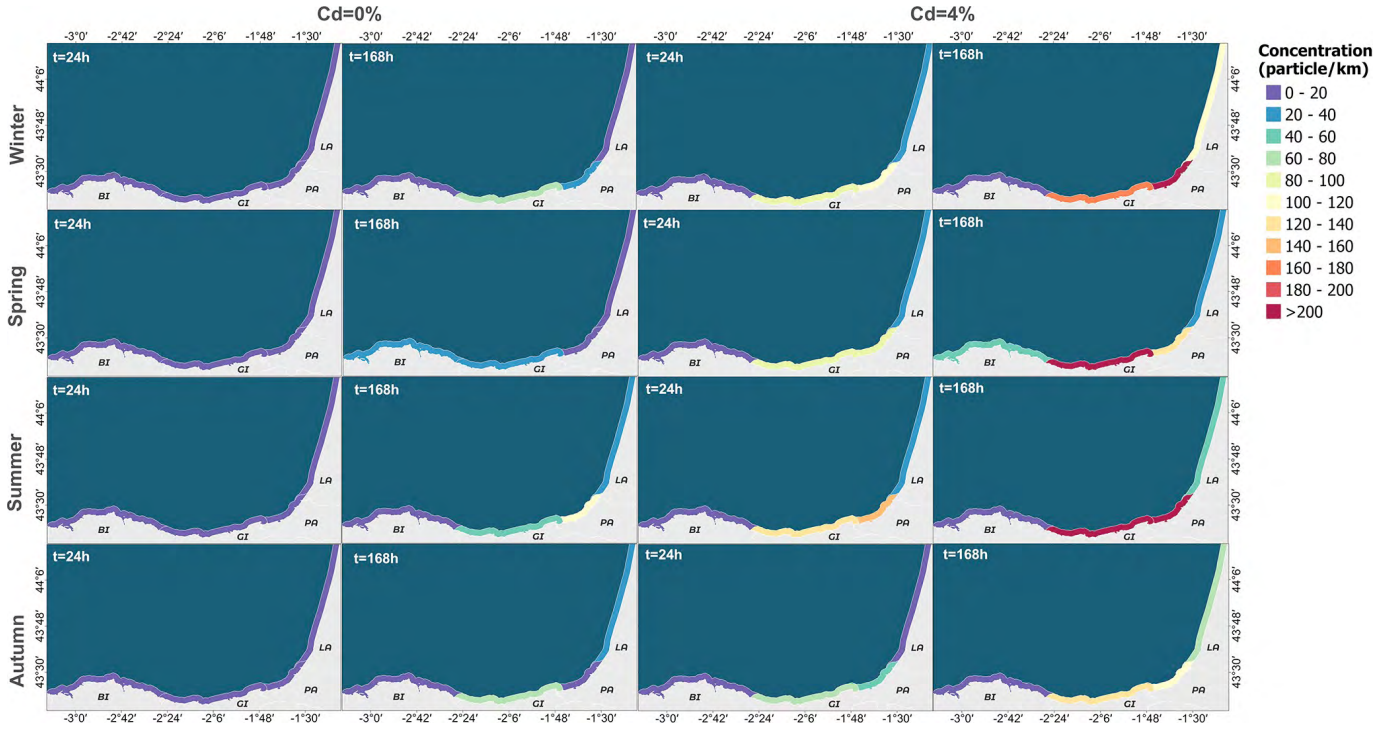


Figure 7: Particle concentration on the Bizkaia, Gipuzkoa, Pyrénées-Atlantiques, and Landes shoreline. The seasonal distribution is shown for $\boldsymbol{C}_d$ = 0% and $\boldsymbol{C}_d$ = 4% after 24 and 168 h of simulation

4.3. *Seasonal trends in floating riverine litter transport and fate*

Particle concentration on the shoreline varied between 0 and 258.46 particles km^{-1} (Fig. 7). Particles parameterized with C_d 4% drifted faster towards the coast, notably during the first 24 h. The highest concentrations (> 200 particles km^{-1}) were recorded during summer in Pyrénées-Atlantiques for C_d 4%, probably due to the seasonal retention patterns within the study area (Appendix D). Although less intensely, C_d = 4% also led to a high particle concentration in Pyrénées-Atlantiques (106.86 particles km^{-1}) and Gipuzkoa (166.1 particles km^{-1}) during winter. The lowest concentrations (0-20 particles km^{-1}) were recorded for C_d 0% after the first 24 h of simulation, particularly during autumn. Overall, Bizkaia was the less impacted region for both windage coefficients (< 40 particles km^{-1}). During summer, over 97% of particles parameterized with C_d 4% beached after 1 week of simulation (Fig. 8). In autumn this value fell to 54%. In contrast, beached particles parameterized with C_d 0% were less abundant by the end of the simulations, particularly during spring with less than 25% of particles trapped in the shoreline.

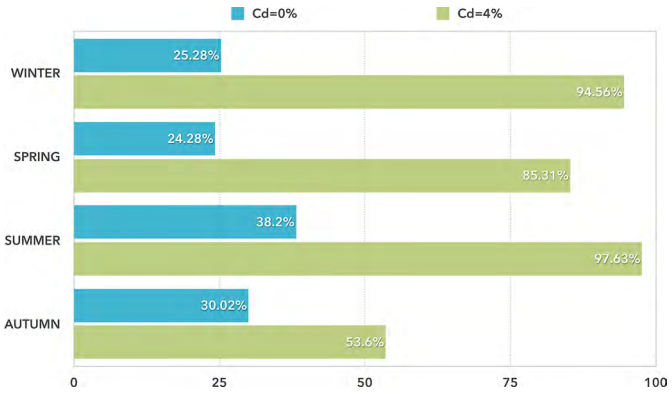


Figure 8: Seasonal numbers of beached particles parameterized with $\boldsymbol{C}_d$ = 0% and $\boldsymbol{C}_d$ = 4% after 168 h of simulation

Overall, the average of particles parameterized with C_d 0% was higher when comparing to C_d 4% (Fig. 9). Particles released in French rivers and parameterized with C_d 0% were less abundant during summer, though this trend was reversed in autumn. For C_d 0%, the number of particles released in the Bidasoa River during summer were the least abundant after 1 week of simulation (< 200 particles on average). The vast majority of particles released in the Urumea River during winter were floating in the study area by the end of the simulations (479 particles on average). Particles parameterized with C_d 4% beached faster during the first 48 h, mainly in summer and for those particles released in the French rivers. During this season, the average number of particles floating in the study area by the end of the simulation ranged between 0 and 250. Similar trends were observed within the same season between rivers, probably influenced by the vicinity of rivers and the spatiotemporal resolution of forcings. Over 40% of the total particles parameterized with C_d 4% and almost 12% parameterized with C_d 0% beached in Gipuzkoa (Fig. 10). During spring, almost 60% of beached particles parameterized with C_d 0% were located Bizkaia. For C_d 0%, particles released during summer in the rivers located in the western area of Gipuzkoa drifted longer distances and reached the Landes shoreline. This trend changed during winter, when the vast majority of particles released in Gipuzkoa rivers beached mainly in Gipuzkoa and Bizkaia. Beached particles parameterized with C_d 0% experienced more seasonal variations derived from the surface current circulation patterns within the SE Bay of Biscay. For C_d 4%, particles beached in Gipuzkoa ranged between 51% in spring and 38% in winter, and Bizkaia was the less affected region despite the season. Overall, all regions were highly affected by rivers within or nearby the region itself.

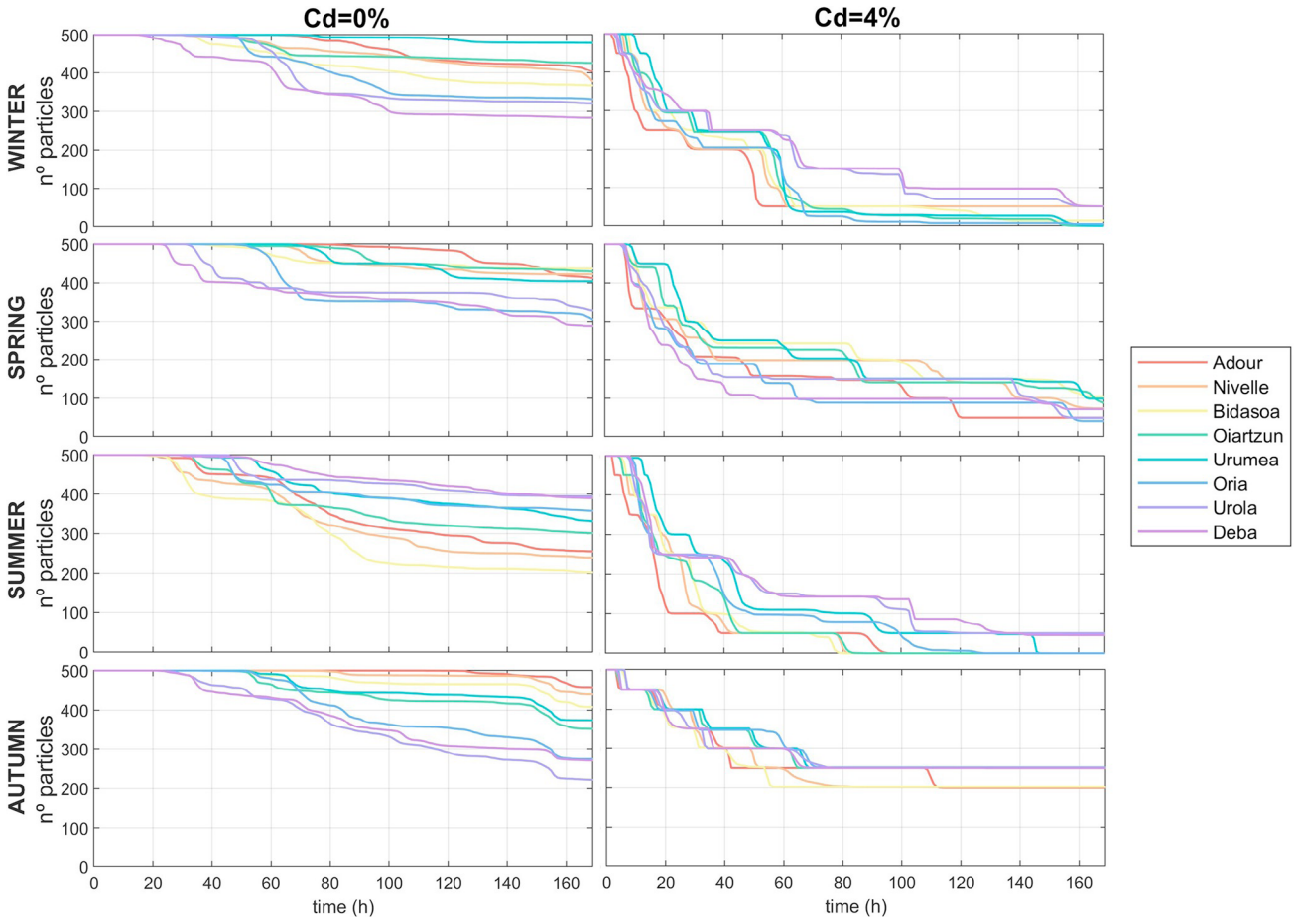


Figure 9: Temporal evolution of the particles parameterized with C_d 0% and C_d 4% throughout the different seasons. The curves represent the average number of particles floating in the water surface by river and for every time step

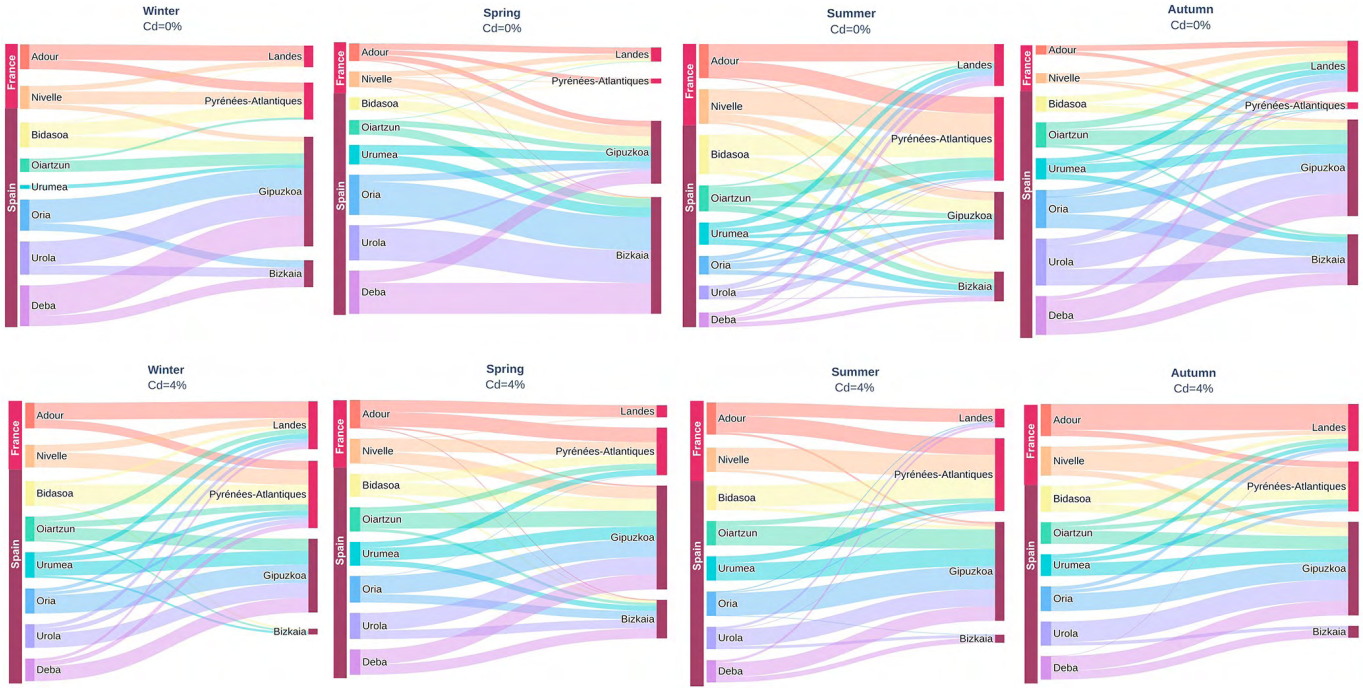


Figure 10: Seasonal analysis of the beached particles parameterized with C_d 0% and C_d 4% per region and river by the end of the simulation period. The nodes of the region correspond to the number of beached particles. The country to which each river belongs —France (Pyrénées-Atlantiques) and Spain (Gipuzkoa)— is shown on the left side of each figure. The width of the node depicts the sum of the beached particles, and the links represent the number of particles beached per river

5. Discussion

5.1. Riverine litter composition

An artisanal net placed at the mouth of the Deba River enabled sampling riverine litter in the study area during spring 2018. Short and narrow rivers prevail in the SE Bay of Biscay, affected by a strong tidal regime and very intense, stationary and persistent storms (Ocio *et al.*, 2015). Studies aiming at reporting the abundance and composition of floating riverine litter in European rivers date back less than 10 years, and they were performed in larger and more abundant rivers than the Deba River. Despite the morphology and hydrological differences, plastic was the predominant material in the Deba River, as in the Siene (Gasperi *et al.*, 2014), Danube (Lechner *et al.*, 2014) or Rhine River (van der Wal *et al.*, 2015). Similarities were also found when comparing the top 10 list of riverine litter items to rivers located in the north-east Atlantic region. Plastic/polystyrene pieces between 2.5 and 50 cm (71.2 %) top the list in terms of the number of items, and their abundance was slightly higher when compared to north-east Atlantic rivers (54.53%) (Bruge *et al.*, 2018; González-Fernández *et al.*, 2018). Lower abundances were observed in the Mediterranean (25.01 %) and the Black Sea (13.74%). Riverine litter items trapped on vegetation or deposited on the riverbank can be degraded by weather conditions (rain, wind, etc.) favouring the fragmentation in plastic pieces before their arrival in the coastal and marine environment (Chamas *et al.*, 2020). The fragmentation can be also influenced by the material and the shape of the litter items (Woods *et al.*, 2021). Differences in plastic/polystyrene pieces between 2.5 and 50 cm abundances can be attributed to a faster fragmentation due to the variations in weather conditions between river basins. However, more detailed analyses on the physical characteristics of litter items (i.e. polymer type) are necessary to fully assess their impact on the occurrence of fragmented plastic pieces. Results are also in line with the ranking list of the top 10 beach litter items across the north-east Atlantic region revealing that single-use plastics (i.e. food containers, bottles, and other packaging) are among the most abundant riverine litter items together with plastic fragments (European Commission, 2018). These results differed from the analysis performed in sea small-scale convergence areas of floating marine

litter ("litter windrows") on the coastal waters of the SE Bay of Biscay, where fishing-related items were the second most abundant sub-category in terms of number after plastic/polystyrene pieces between 2.5 and 50 cm (Ruiz *et al.*, 2020). Substantial differences also exist between riverine litter sampled in the Deba River and floating marine litter assessed by visual observation from research vessels in open waters of the Bay of Biscay (Ruiz *et al.*, 2022a). Differences might be related to the monitoring method and, also, to the size of the items, since small items, such as plastic pieces, can be overlooked by the observer when the visual counting method is applied, contrary to riverine litter samplings for later analysis in the lab. Overall, riverine litter data acquisition is mainly focused on the floating fraction, and the litter loads under the surface water are often ignored. Increasing the quantity of rivers sampled, the frequency, and the riverine water compartments is necessary to establish the composition and trends of riverine litter in the SE Bay of Biscay.

5.2. *Wind drag estimation*

One of the largest uncertainties for predicting floating riverine and marine litter behaviour is the proper quantification of a wind drag coefficient. Wind drag estimations conducted so far for floating marine litter items range between 0% and 6% (Ko *et al.*, 2020; Critchell and Lambrechts, 2016; Neumann *et al.*, 2014) with an upper limit of 10% (Yoon *et al.*, 2010). However, only a few of them have been validated using observational data (Maximenko *et al.*, 2018; Callies *et al.*, 2017). In this study, data provided by low-cost buoys combined with surface current measurements by HF radar were used as a proxy for modelling the drift of floating litter objects with similar buoy characteristics (density, size, and shape). Results demonstrated that C_d 4% was the optimal wind drag coefficient for accurately represent the pathways of the low-cost buoys in the study area. This value can be consistent with the estimations of the partially emerged *Physalia physalis* for the Bay of Biscay (Ferrer and Pastor, 2017), but it is almost 3 times higher than the maximum wind drag coefficient reported in the area by Pereiro *et al.* (2018). This can be explained by the fact that buoys used in the experiment remained submerged beneath the sea surface and were less exposed to the wind effect. The estimated wind drag coefficient was also greater than C_d 3 % observed for the Prestige oil spill accident (Abascal *et al.*, 2009; Marta-Almeida *et al.*, 2013). Indeed, oil spill studies refer to a range of wind drag coefficient between 2.5% to 4.4% of the wind speed, with a mean value of 3%-3.5% (e.g. ASCE, 1996; Reed *et al.*, 1994). Object characteristics may change over time due to the exposure to wind, waves, UV radiation, seawater, and the attachment of organic material (Kooi *et al.*, 2017; Min *et al.*, 2020). Objects become breakable, and biofouling increases their density, overcoming the positive buoyancy and impacting their trajectory. Investigations so far pinpointed longer timescales (weeks to months and longer) than considered in this study (days) for a significant change on the behaviour of floating objects (Ryan, 2015; Fazey and Ryan, 2016). Consequently, physical variations in the buoy properties were not accounted for the wind drag estimation. The separation distance between observed and modelled trajectories has been commonly used to evaluate the skill of particle-tracking models (Callies *et al.*, 2017; Haza *et al.*, 2019; Aksamit *et al.*, 2020; Abascal *et al.*, 2012). In this study, the purpose was not to evaluate the model accuracy but to estimate the wind drag coefficient for the low-cost buoys. However, the novel approach proposed by Révelard *et al.* (2021) may be of particular interest for future experiments oriented towards assessing the wind drag coefficient of highly buoyant items drifting for short time periods in the coastal area.

5.3. *Seasonal riverine litter distribution by region*

It is broadly accepted that the SE Bay of Biscay is polluted with floating marine litter discarded or lost in the marine and coastal area but also with litter originating inland and transported via rivers and run-off. However, detailed studies on riverine litter contribution are still scarce, and modelling efforts combining observations and physical parameterizations of floating litter properties are non-existent. This study shows that the exposure to the wind effect largely controls the transport and coastal accumulation of floating marine litter in the SE Bay of Biscay, with concentrations varying between regions and over time. Concentrations in Pyrénées-Atlantiques and Gipuzkoa differed widely from the other studied regions. Indeed, the highest

concentrations occurred in both regions during summer for low- (100-120 particle km^{-1}) and high-buoyancy items (> 200 particles km^{-1}). A higher number of particles beached in Gipuzkoa during summer when compared to Pyrénées-Atlantiques, but concentrations were lower since the Basque shoreline is longer. The pathways and fate of low-buoyancy items reflect the seasonal surface water circulation patterns in the SE Bay of Biscay. Results are in line with findings provided by Declerck *et al.* (2019) who pinpointed a higher coastal retention in the area during spring and summer. Weakly buoyant objects remained floating at the coastal waters and highly buoyant objects tended to beach remarkably faster as reported in literature by Rodríguez-Díaz *et al.* (2020). However, long-term data collected by *in situ* observations of beached litter across the different regions are necessary to validate the large seasonal variations and to assess the reliability of concentration levels for addressing riverine litter issue in priority regions with heavily polluted coastlines.

5.4. *Rivers as key vectors of riverine litter*

The interpretation of the spatial and temporal riverine litter distribution by river can be challenging since riverine litter fluxes in the study area are highly uncertain. In the study area, two major assumptions were made regarding the river systems: (1) the same river discharge for all rivers and (2) the same river discharge for all seasons. This means that the same amounts of riverine litter were allocated for every river regardless of the differences in the width and depth and the seasonal flow variations. Since each river basin has its own particularities, future modelling approaches should be adapted to the morphology and hydrological conditions of the catchment area. Other drivers, such as the land use or population density, can be a determining factor for the amount of mismanaged litter that could contribute to riverine litter fluxes (Schmidt *et al.*, 2017; Schuyler *et al.*, 2021). It is also necessary to further investigate if higher river flows in the area are directly related to an increased discharge of riverine litter since analysis already performed in different river basins shows contradicting relations between the occurrence of riverine litter and river fluxes (van Emmerik and Schwarz, 2020). Along with the complex nature of qualifying riverine litter fluxes, litter behaviour in the coastal area of the SE Bay of Biscay is still in its early stage, and much has yet to be revealed. Particular attention should be paid to Pyrénées-Atlantiques and Gipuzkoa, as the main impacted regions in the studied area. Rivers in the study area are mainly located in Gipuzkoa, which favours the accumulation of floating litter in this region regardless of the season. Regional coordination should be reinforced due to the transboundary movement of floating riverine litter in the study area and reasonable efforts oriented towards retaining or removing riverine litter as clean-up measures in the riverbanks should be investigated to avoid litter being transported to the coastal and marine environment.

5.5. *Model limitations*

The interaction between floating litter and the shoreline is highly complex and relies on many processes including waves and tides. Indeed, waves and tides can constrain coastal accumulation since they can resuspend and transport litter back into the ocean (Brennan *et al.*, 2018; Compa *et al.*, 2022). The geomorphology can also affect the retention of litter washing ashore. Sandy beaches tend to be more efficient at trapping and accumulating litter than rocky areas, which favour litter fragmentation (Robbe *et al.*, 2021; Weideman *et al.*, 2020). How these processes contribute to the actual beaching is unknown, and they cannot be resolved yet at a suitable resolution (Melvin *et al.*, 2021). In this study, particles were released in open waters, and once they reached the shoreline, they were classified as beached. The tidal effect and the wave-induced Stokes drift were not accounted for to avoid introducing more uncertainties. However, further field and laboratory experiments to better understand how these processes influence floating litter behaviour on the coastline are recommend. For future research, it is also important to consider exploring the effect of the type of shore-line on coastal accumulation. In this study, a constant diffusion coefficient of 1 m^2 s^{-1} was regarded as a pragmatic choice based on previous modelling work for floating marine litter. However, more field measurements are necessary to accurately assess the influence of the diffusion process on the transport of floating marine litter.

6. Conclusions

The SE Bay of Biscay has been described by global and regional models as an accumulation zone for floating marine litter. However, detailed studies on floating riverine litter behaviour once items arrive in open waters are still scarce. Based on HF radar current observations and a wind dataset, this contribution tries to fill this gap by providing insights into how low- and high-buoyancy litter released by several rivers of the SE Bay of Biscay may affect the nearby regions seasonally in terms of concentration and beaching. Analysis of riverine litter samples collected by a barrier placed in the study area showed that low-buoyancy objects were predominant, although highly buoyant objects were also relevant in terms of weight. Simulations for assessing the seasonal trends of floating riverine litter transport and fate were performed with the Lagrangian model TESEO. To properly integrate the differences in litter buoyancy, simulations were parameterized with a wind drag coefficient for low- and high-buoyancy items. The wind drag for highly buoyant items was estimated by comparing the observed and the modelled positions of four drifters. The developed low-cost buoys proved to be suitable to provide real-time trajectories of highly buoyant objects exposed to wind. However, drifters with different characteristics should be used in future studies to account for the windage effect on different types of items. The transport and fate of both high- and low-buoyancy items released by rivers was calculated by season. Highly buoyant items rapidly beached (in less than 48 h), particularly in summer and winter; in contrast, despite the season over two-thirds of low-buoyancy items remained floating after 1 week of being released. This highlights the discrepancy between the behaviour for low- and high-buoyancy objects and the importance of parameterizing the windage effect in order to accurately predict riverine litter accumulation in the coastal area of the SE Bay of Biscay. Beached particles were mainly found in Gipuzkoa regardless of the season and the wind drag coefficient. Overall, the less affected region was Bizkaia with the exception of a spring period for low-buoyancy items. Despite the season, most of the riverine litter remained in the study area and rivers polluted the regions within the river basin or surrounding it. Investigating what beaches are most likely to accumulate large quantities and the contribution per river can provide relevant input to response operations after storm events in the short to medium term and can also support the identification of priority rivers for a monitoring programme, assisting adapted intervention of riverine pollution regionally in the future.

Code and data availability

Code and data used to conduct this study are available upon request by contacting the corresponding authors.

Video supplement

Animations of the surface currents, winds, and Lagrangian simulations area available for the study period 2009-2021 (https://doi.org/10.5446/s_1355, Ruiz *et al.*, 2022b).

Author contributions

IR performed the investigation, the data analysis, and the visualization assets and wrote the original paper. AJA contributed to the conceptualization of the investigation, provided the software, and reviewed and edited the paper. OCB and AR contributed to the conceptualization of the investigation and supervised, reviewed, and edited the paper. All authors contributed to refining the paper for submission.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

We are grateful to the Emergencies and Meteorology Directorate —Security department— Basque Government for public data provision from the Basque Operational Oceanography System EuskOOS. This study has been conducted using EU Copernicus Marine Service information. We thank Luis Ferrer for sharing his valuable knowledge on the custom-built drifters. We thank Cristina Barreau, Antoine Bruge, Igor Granado, Théo Destang, Alix McDaid, and Jon Andonegi for their support with the release of the drifters. We thank the citizens who collected and reported the drifters' arrival at the Basque and French coasts. This paper is contribution no. 1134 from AZTI, Marine Research, Basque Research and Technology Alliance (BRTA). This paper is part of the PhD research of Irene Ruiz, supervised by Oihane C. Basurko and Anna Rubio.

Financial support

This research has been partially funded through the EU's LIFE Program (LIFE LEMA project, grant agreement no. LIFE15 ENV/ES/000252) and by the EU's H2020 Program (JERICO-S3 project, grant agreement no. 871153).

Review statement

This paper was edited by Erik van Sebille and reviewed by two anonymous referees.

References

Abascal, A., Castanedo, S., Gutierrez, A. D., Comerma, E., Medina, R., and Losada, I. J.: Teseo, an operational system for simulating oil spills trajectories and fate processes, Proc. Int. Offshore Polar Eng. Conf., 1751-1758, ISBN 978-1-880653-68-5, 2007.

Abascal, A. J., Castanedo, S., Mendez, F. J., Medina, R., and Losada, I. J.: Calibration of a Lagrangian Transport Model Using Drifting Buoys Deployed during the *Prestige* Oil Spill, J. Coastal Res., 11, 80-90, 2009.

Abascal, A. J., Castanedo, S., Fernández, V., and Medina, R.: Backtracking drifting objects using surface currents from high-frequency (HF) radar technology, Ocean Dynam., 62, 1073-1089, https://doi.org/10.1007/s10236-012-0546-4, 2012.

Abascal, A. J., Castanedo, S., Núñez, P., Mellor, A., Clements, A., Pérez, B., Cárdenas, M., Chiri, H., and Medina, R.: A high-resolution operational forecast system for oil spill response in Belfast Lough, Mar. Pollut. Bull., 114, 302-314, https://doi.org/10.1016/j.marpolbul.2016.09.042, 2017a.

Abascal, A. J., Sanchez, J., Chiri, H., Ferrer, M. I., Cárdenas, M., Gallego, A., Castanedo, S., Medina, R., Alonso-Martirena, A., Berx, B., Turrell, W. R., and Hughes, S. L.: Operational oil spill trajectory modelling using HF radar currents: A northwest European continental shelf case study, Mar. Pollut. Bull., 119, 336-350, https://doi.org/10.1016/j.marpolbul.2017.04.010, 2017b.

Aksamit, N. O., Sapsis, T., and Haller, G.: Machine-Learning Mesoscale and Submesoscale Surface Dynamics from Lagrangian Ocean Drifter Trajectories, J. Phys. Oceanogr., 50, 1179-1196, https://doi.org/10.1175/jpo-d-19-0238.1, 2020.

Al-Zawaidah, H., Ravazzolo, D., and Friedrich, H.: Macroplastics in rivers: present knowledge, issues and challenges, Environ. Sci.-Proc. Imp., 23, 535-552, https://doi.org/10.1039/d0em00517g, 2021.

Allshouse, M. R., Ivey, G. N., Lowe, R. J., Jones, N. L., Beegle-Krause, C. J., Xu, J., and Peacock, T.: Impact of windage on ocean surface Lagrangian coherent structures, Environ. Fluid Mech., 17, 473-483, https://doi.org/10.1007/s10652-016-9499-3, 2017.

ASCE: State-of-the-Art Review of Modeling Transport and Fate of Oil Spills, J. Hydraul. Eng., 122, 594-609, https://doi.org/10.1061/(ASCE)0733-94 29(1996)122:11(594), 1996.

Blettler, M. C. M., Abrial, E., Khan, F. R., Sivri, N., and Espinola, L. A.: Freshwater plastic pollution: Recognizing research bi-ases and identifying knowledge gaps, Water Res., 143, 416-424, https://doi.org/10.1016/j.watres.2018.06.015, 2018.

Bourillet, J. F., Zaragosi, S., and Mulder, T.: The French Atlantic margin and deep-sea submarine systems, Geo-Mar. Lett., 26, 311-315, https://doi.org/10.1007/s00367-006-0042-2, 2006.

Brennan, E., Wilcox, C., and Hardesty, B. D.: Connecting flux, deposition and resuspension in coastal debris surveys, Sci. Total Environ., 644, 1019-1026, https://doi.org/10.1016/J.SCITOTENV.2018.06.352, 2018.

Bruge, A., Barreau, C., Carlot, J., Collin, H., Moreno, C., and Maison, P.: Monitoring Litter Inputs from the Adour River (South-west France) to the Marine Environment, Mar. Sci. Eng., 6, 24, https://doi.org/10.3390/jmse6010024, 2018.

C3S: ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate, https://cds.climate.copernicus.eu/cdsapp#!/home, last access: 11 November 2019.

Callies, U., Groll, N., Horstmann, J., Kapitza, H., Klein, H., Maßmann, S., and Schwichtenberg, F.: Surface drifters in the German Bight: model validation considering windage and Stokes drift, Ocean Sci., 13, 799-827, https://doi.org/10.5194/os-13-799-2017, 2017.

Carlson, D. F., Pavalko, W. J., Petersen, D., Olsen, M., and Hass, A. E.: Maker Buoy Variants for Water Level Monitoring and Tracking Drifting Objects in Remote Areas of Greenland, 20, 1254, https://doi.org/10.3390/s20051254, 2020.

Carson, H. S., Lamson, M. R., Nakashima, D., Toloumu, D., Hafner, J., Maximenko, N., and McDermid, K. J.: Tracking the sources and sinks of local marine debris in Hawai'i, Mar. Environ. Res., 84, 76-83, 2013.

Chamas, A., Moon, H., Zheng, J., Qiu, Y., Tabassum, T., Hee Jang, J., Abu-Omar, M., Scott, S. L., and Suh, S.: Degradation Rates of Plastics in the Environment, ACS Sustain. Chem. Eng, 8, 3511, https://doi.org/10.1021/acssuschemeng.9b06635, 2020.

Charria, G., Lazure, P., Le Cann, B., Serpette, A., Reverdin, G., Louazel, S., Batifoulier, F., Dumas, F., Pichon, A., and Morel, Y.: Surface layer circulation derived from Lagrangian drifters in the Bay of Biscay, J. Marine Syst., 109, S60-S76, https://doi.org/10.1016/j.jmarsys.2011.09.015, 2013.

Chiri, H., Abascal, A. J., Castanedo, S., and Medina, R.: Mid-long term oil spill forecast based on logistic regression modelling of met-ocean forcings, Mar. Pollut. Bull., 146, 962-976, https://doi.org/10.1016/j.marpolbul.2019.07.053, 2019.

Chiri, H., Abascal, A. J., and Castanedo, S.: Deep oil spill hazard assessment based on spatio-temporal met-ocean patterns, Mar. Pollut. Bull., 154, 111123, https://doi.org/10.1016/J.MARPOLBUL.2020.111123, 2020.

Compa, M., Alomar, C., Morató, M., Álvarez, E., and Deudero, S.: Spatial distribution of macro- and micro-litter items along rocky and sandy beaches of a Marine Protected Area in the western Mediterranean Sea, Mar. Pollut. Bull., 178, 113520, https://doi.org/10.1016/J.MARPOLBUL.2022.113520, 2022.

Critchell, K. and Lambrechts, J.: Modelling accumulation of marine plastics in the coastal zone; what are the dominant physical processes?, Estuar. Coast. Shelf S., 171, 111, https://doi.org/10.1016/j.ecss.2016.01.036, 2016.

D'Asaro, E. A., Carlson, D. F., Chamecki, M., Harcourt, R. R., Haus, B. K., Fox-Kemper, B., Molemaker, M. J., Poje, A. C., and Yang, D.: Advances in Observing and Understanding Small-Scale Open Ocean Circulation During the Gulf of Mexico Research Initiative Era, Front. Mar. Sci., 7, 349, https://doi.org/10.3389/fmars.2020.00349, 2020.

Declerck, A., Delpey, M., Rubio, A., Ferrer, L., Basurko, O. C., Mader, J., and Louzao, M.: Transport of floating marine litter in the coastal area of the south-eastern Bay of Biscay: A Lagrangian approach using modelling and observations, J. Oper. Oceanogr., 12, S111-S125, https://doi.org/10.1080/1755876x.2019.1611708, 2019.

Delpey, M., Declerck, A., Epelde, I., Voirand, T., Manso-Navarte, I., Mader, J., Rubio, A., and Caballero, A.: Tracking floating marine litter in the coastal area by combining operational ocean modelling and remote observation systems., EGU General Assembly 2021, online, 19-30 Apr 2021, EGU21-11465, https://doi.org/10.5194/egusphere-egu21-11465, 2021.

Duhec, A. V., Jeanne, R. F., Maximenko, N., and Hafner, J.: Composition and potential origin of marine debris stranded in the Western Indian Ocean on remote Alphonse Island, Seychelles, Mar. Pollut. Bull., 96, 76-86, https://doi.org/10.1016/j.marpolbul.2015.05.042, 2015.

European Commission (Joint Research Centre, Addamo, A., Laroche, P., Hanke, G.): Top marine beach litter items in Europe, European Commission, Joint Research Centre, Publications Office of the European Union, https://doi.org/10.2760/496717, 2018.

Eurostat: Population density by NUTS 3 region, http://data.europa.eu/88u/dataset/gngfvpqmfu5n6akvxqkpw (last access: 21 February 2020), 2019.

Fazey, F. M. C. and Ryan, P. G.: Biofouling on buoyant marine plastics: An experimental study into the effect of size on surface longevity, Environ. Pollut., 210, 354-360, https://doi.org/10.1016/J.ENVPOL.2016.01.026, 2016.

Ferrer, L. and Pastor, A.: The Portuguese man-of-war: Gone with the wind, Reg. Stud. Mar. Sci., 14, 53-62, https://doi.org/10.1016/j.rsma.2017.05.004, 2017.

Gasperi, J., Dris, R., Bonin, T., Rocher, V., and Tassin, B.: Assessment of floating plastic debris in surface water along the Seine River, Environ. Pollut., 195, 163-166, https://doi.org/10.1016/j.envpol.2014.09.001, 2014.

González, M., Valencia, V., Mader, J., Fontán, A., Uriarte, A., and Caballero, A.: Operational Coastal Systems in the Basque Country Region: Modelling and Observations, Seventeenth International Offshore and Polar Engineering Conference, Lisbon, Portugal, 1 July 2007, Proc. Int. Offshore Polar Eng. Conf., ISOPE-I-07-074, 2007.

González-Fernández, D. and Hanke, G.: Toward a Harmonized Approach for Monitoring of Riverine Floating Macro Litter Inputs to the Marine Environment, Front. Mar. Sci., 4, 86, https://doi.org/10.3389/fmars.2017.00086, 2017.

González-Fernández, D., Hanke, G., Kideys, A., Navarrao Ortega, A., Sanchez Vidal, A., Bruge, A., Öztürk, B., Palma, C., Santelli, C., Duijsings, D., Barcelo, D., Dimitiriu, E., Rojo-Nieto, E., Ferreira, F., Bessa, F., Suaria, G., Siedlewicz, G., Castro Jimenez, J., Germano, J., Pereira De Brito, J., Rigueira, J., Pazdro, K., Cabrera, M., Pogojeva, M., Köck Schulmeyer, M., Constant, M., Canals Artigas, M., Paraboschi, M., Tourgeli, M., Machitadze, N., Ratola, N., Savenko, O., Kerherve, P., Sempere, R., Bakiu, R., Crosti, R., Schoeneich-Argent, R., Landry Levesque, S., Agostinho, T., Segal, Y., and Galletti, Y.: Floating Macro Litter in European Rivers – Top Items, EUR 29383 EN, Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/316058, 2018.

González-Fernández, D., Cózar, A., Hanke, G., Viejo, J., Morales-Caselles, C., Bakiu, R., Barceló, D., Bessa, F., Bruge, A., Cabrera, M., Castro-Jiménez, J., Constant, M., Crosti, R., Galletti, Y., Kideys, A. E., Machitadze, N., Pereira de Brito, J., Pogojeva, M., Ratola, N., Rigueira, J., Rojo-Nieto, E., Savenko, O., Schöneich-Argent, R. I., Siedlewicz, G., Suaria, G., and Tourgeli, M.: Floating macrolitter leaked from Europe into the ocean, Nat. Sustain., 4, 474-483, https://doi.org/10.1038/s41893-021-00722-6, 2021.

Haza, A. C., Paldor, N., Özgökmen, T. M., Curcic, M., Chen, S. S., and Jacobs, G.: Wind-Based Estimations of Ocean Surface Currents From Massive Clusters of Drifters in the Gulf of Mexico, J. Geophys. Res.-Oceans, 124, 5844-5869, https://doi.org/10.1029/2018JC014813, 2019.

Hernández-Carrasco, I., Solabarrieta, L., Rubio, A., Esnaola, G., Reyes, E., and Orfila, A.: Impact of HF radar current gap-filling methodologies on the Lagrangian assessment of coastal dynamics, Ocean Sci., 14, 827-847, https://doi.org/10.5194/os-14-827-2018, 2018.

Hoenner, X., Barlian, E., Ernawati, T., Hardesty, B. D., Kembaren, D. D., Mous, P. J., Sadiyah, L., Satria, F., and Wilcox, C.: Using anti-theft tracking devices to infer fishing vessel activity at sea, Fish. Res., 249, 106230, https://doi.org/10.1016/j.fishres.2022.106230, 2022.

Hunter, J. R., Craig, P. D., and Phillips, H. E.: On the use of random walk models with spatially variable diffusivity, J. Comput. Phys., 106, 366-376, https://doi.org/10.1016/S0021-9991(83)71114-9, 1993.

Joint Research Centre (Institute for Environment and Sustainability): Guidance on monitoring of marine litter in European seas, Publications Office, https://doi.org/10.2788/99816, 2014.

Kaplan, D. M. and Lekien, F.: Spatial interpolation and filtering of surface current data based on open-boundary modal analysis, J. Geophys. Res.-Oceans, 112, C12007, https://doi.org/10.1029/2006JC003984, 2007.

Karagiorgos, J., Vervatis, V., and Sofianos, S.: The Impact of Tides on the Bay of Biscay Dynamics, J. Mar. Sci. Eng., 8, 617, https://doi.org/10.3390/jmse8080617, 2020.

Ko, C.-Y., Hsin, Y.-C., and Jeng, M.-S.: Global distribution and clean-up opportunities for macro ocean litter: a quarter century of accumulation dynamics under windage effects, Environ. Res. Lett., 15, 104063, https://doi.org/10.1088/1748-9326/abae29, 2020.

Kooi, M., Nes, E. H. van, Scheffer, M., and Koelmans, A. A.: Ups and Downs in the Ocean: Effects of Biofouling on Vertical Transport of Microplastics, Environ. Sci. Technol., 51, 7963-7971, https://doi.org/10.1021/acs.est.6b04702, 2017.

Lavin, A., Valdés, L., Sanchez, F., Abaunza, P., Forest, A., Boucher, J., Lazure, P., and Jegou, A.-M.: The Bay of Biscay: the encountering of the Ocean and the Shelf (18b, E), edited by: Robinson, A. and Brink, K., Combridge, MA, Harvard University Press, http://hdl.handle.net/10508/3128 (last access: 25 November 2022) 2006.

Le Boyer, A., Charria, G., Le Cann, B., Lazure, P., and Marié, L.: Circulation on the shelf and the upper slope of the Bay of Biscay, Cont. Shelf Res., 55, 97-107, 2013.

Lebreton, L. C. M., Greer, S. D., and Borrero, J. C.: Numerical modelling of floating debris in the world's oceans, Mar. Pollut. Bull., 64, 653-661, 2012.

Lebreton, L., Egger, M., and Slat, B.: A global mass budget for positively buoyant macroplastic debris in the ocean, Sci. Rep.-UK, 9, 12922, https://doi.org/10.1038/s41598-019-49413-5, 2019.

Lebreton, L. C. M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., and Reisser, J.: River plastic emissions to the world's oceans, Nat. Commun., 8, 15611, https://doi.org/10.1038/ncomms15611, 2017.

Lechner, A., Keckeis, H., Lumesberger-Loisl, F., Zens, B., Krusch, R., Tritthart, M., Glas, M., and Schludermann, E.: The Danube so colourful: A potpourri of plastic litter outnumbers fish larvae in Europe's second largest river, Environ. Pollut., 188, 177-181, https://doi.org/10.1016/J.ENVPOL.2014.02.006, 2014.

Maclean, K., Weideman, E. A., Perold, V., and Ryan, P. G.: Buoyancy affects stranding rate and dispersal distance of floating litter entering the sea from river mouths, Mar. Pollut. Bull., 173, 113028, https://doi.org/10.1016/j.marpolbul.2021.113028, 2021. Mai, L., Sun, X., Xia, L.-L., Bao, L.-J., Liu, L.-Y., and Zeng, E. Y.: Global Riverine Plastic Outflows, Environ. Sci. Technol., 54, 10049-10056, https://doi.org/10.1021/acs.est.0c02273, 2020.

Maier-Reimer, E. and Sündermann, J.: On tracer methods in computational hydrodynamics, in: Engineering applications of computational hydraulics, edited by: Abbott, M. B., 198-216, https://hdl.handle.net/21.11116/0000-0004-F582-4 (last access: 25 November 2022), 1982.

Manso-Narvarte, I., Caballero, A., Rubio, A., Dufau, C., and Birol, F.: Joint analysis of coastal altimetry and high-frequency (HF) radar data: observability of seasonal and mesoscale ocean dynamics in the Bay of Biscay, Ocean Sci., 14, 1265-1281, https://doi.org/10.5194/os-14-1265-2018, 2018.

Manso-Narvarte, I., Rubio, A., Jordà, G., Carpenter, J., Merckelbach, L., and Caballero, A.: Three-Dimensional Characterization of a Coastal Mode-Water Eddy from Multiplatform Observations and a Data Reconstruction Method, Remote Sens., 13, 674, https://doi.org/10.3390/rs13040674, 2021.

Mantovani, C., Corgnati, L., Horstmann, J., Rubio, A., Reyes, E., Quentin, C., Cosoli, S., Asensio, J. L., Hermes, J., Mader, J., and Griffa, A.: Best Practices on High Frequency Radar Deployment and Operation for Ocean Current Measurement, Front. Mar. Sci., 7, 210, https://doi.org/10.3389/fmars.2020.00210, 2020.

Margenat, H., Ruiz-Orejón, L. F., Cornejo, D., Martí, E., Vila, A., Le Roux, G., Hansson, S., and Guasch, H.: Guidelines of field-tested procedures and methods for monitoring plastic litter in Mountain riverine systems, European Commission, https://doi.org/10.20350/digitalCSIC/13880, 2021.

Marta-Almeida, M., Ruiz-Villarreal, M., Pereira, J., Otero, P., Cirano, M., Zhang, X., and Hetland, R. D.: Efficient tools for marine operational forecast and oil spill tracking, Mar. Pollut. Bull., 71, 139-151, https://doi.org/10.1016/J.MARPOLBUL.2013.03.022, 2013.

Martínez Fernández, A., Redondo Caride, W., Alonso Pérez, F., Piedracoba Varela, S., Lorente Jiménez, P., Montero Vilar, P., Torres López, S., Fernández Baladrón, A., Varela Benvenuto, R. A., and Velo Lanchas, A.: SPOT and GPRS drifting buoys for HF Radar calibration, Instrum. Viewp., Universidad Politécnica de Cataluña, 48-49, http://hdl.handle.net/10261/259925 (last access: 25 November 2022), 2021.

Maximenko, N., Hafner, J., Kamachi, M., and MacFadyen, A.: Numerical simulations of debris drift from the Great Japan Tsunami of 2011 and their verification with observational reports, Mar. Pollut. Bull., 132, 5-25, https://doi.org/10.1016/j.marpolbul.2018.03.056, 2018.

Mazarrasa, I., Puente, A., Núñez, P., García, A., Abascal, A. J., and Juanes, J. A.: Assessing the risk of marine litter accumulation in estuarine habitats, Mar. Pollut. Bull., 144, 117-128, https://doi.org/10.1016/j.marpolbul.2019.04.060, 2019.

Meijer, L. J. J., van Emmerik, T., van der Ent, R., Schmidt, C., and Lebreton, L.: More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean, Sci. Adv., 7, eaaz5803, https://doi.org/10.1126/sciadv.aaz5803, 2021.

Melvin, J., Bury, M., Ammendolia, J., Charles, M., and Liboiron, M.: Critical Gaps in Shoreline Plastics Pollution Research, Front. Mar. Sci., 8, 689108, https://doi.org/10.3389/fmars.2021.689108, 2021.

Meyerjürgens, J., Badewien, T. H., Garaba, S. P., Wolff, J.-O., and Zielinski, O.: A State-of-the-Art Compact Surface Drifter Reveals Pathways of Floating Marine Litter in the German Bight, Front. Mar. Sci., 6, 58, https://doi.org/10.3389/fmars.2019.00058, 2019.

Min, K., Cuiffi, J. D., and Mathers, R. T.: Ranking environmental degradation trends of plastic marine debris based on physical properties and molecular structure, Nat. Commun., 11, 727, https://doi.org/10.1038/s41467-020-14538-z, 2020.

Neumann, D., Callies, U., and Matthies, M.: Marine litter ensemble transport simulations in the southern North Sea, Mar. Pollut. Bull., 86, 219-228, https://doi.org/10.1016/j.marpolbul.2014.07.016, 2014.

Novelli, G., Guigand, C. M., Cousin, C., Ryan, E. H., Laxague, N. J. M., Dai, H., Haus, B. K., and Özgökmen, T. M.: A Biodegradable Surface Drifter for Ocean Sampling on a Massive Scale, J. Atmos. Ocean. Tech., 34, 2509-2532, https://doi.org/10.1175/jtech-d-17-0055.1, 2017.

Novelli, G., Guigand, C. M., and Özgökmen, T. M.: Technological Advances in Drifters for Oil Transport Studies, Mar. Technol. Soc. J., 52, 53-61, https://doi.org/10.4031/mtsj.52.6.9, 2018.

Núñez, P., García, A., Mazarrasa, I., Juanes, J. A., Abascal, A. J., Méndez, F., Castanedo, S., and Medina, R.: A methodology to assess the probability of marine litter accumulation in estuaries, Mar. Pollut. Bull., 144, 309-324, https://doi.org/10.1016/j.marpolbul.2019.04.077, 2019.

Ocio, D., Stocker, C., Eraso, A., and Cowpertwait, P.: Regionalized extreme flows by means of stochastic storm generation coupled with a distributed hydrological model. The case of the Basque Country, Proc. 36th IAHR World Congr., 28 June-3 July 2015, 1-13, https://doi.org/10.13140/RG.2.1.3924.4889, 2015.

Pereiro, D., Souto, C., and Gago, J.: Calibration of a marine floating litter transport model, J. Oper. Oceanogr., 11, 125-133, https://doi.org/10.1080/1755876x.2018.1470892, 2018.

Pereiro, D., Souto, C., and Gago, J.: Dynamics of floating marine debris in the northern Iberian waters: A model approach, J. Sea Res., 144, 57-66, https://doi.org/10.1016/j.seares.2018.11.007, 2019.

Porter, M., Inall, M. E., Green, J. A. M., Simpson, J. H., Dale, A. C., and Miller, P. I.: Drifter observations in the summer time Bay of Biscay slope current, J. Marine Syst., 157, 65-74, 2016.

Puillat, I., Lazure, P., Anne-marie, J., Lampert, L., and Miller, P.: Mesoscale hydrological variability induced by northwesterly wind on the French continental shelf of the Bay of Biscay, Sci. Mar., 70, 15-26, 2006.

Putman, N. F., Lumpkin, R., Olascoaga, M. J., Trinanes, J., and Goni, G. J.: Improving transport predictions of pelagic Sargassum, J. Exp. Mar. Biol. Ecol., 529, 151398, https://doi.org/10.1016/j.jembe.2020.151398, 2020.

Reed, M., Turner, C., and Odulo, A.: The role of wind and emulsification in modelling oil spill and surface drifter trajectories, Spill Sci. Technol. B., 1, 143-157, https://doi.org/10.1016/1353-2561(94)90022-1, 1994.

Révelard, A., Reyes, E., Mourre, B., Hernández-Carrasco, I., Rubio, A., Lorente, P., Fernández, C. D. L., Mader, J., Álvarez-Fanjul, E., and Tintoré, J.: Sensitivity of Skill Score Metric to Validate Lagrangian Simulations in Coastal Areas: Recommendations for Search and Rescue Applications, Front. Mar. Sci., 8, 630388, https://doi.org/10.3389/fmars.2021.630388, 2021.

Rizal, A., Gautama, B. G., Pranowo, W. S., Farhan, A. R., Siong, K., Harjono, M. R., Voisin, J. B., Maes, C., Dobler, D., Berlianty, D., Priyono, B., Dufau, C., Lucas, M., Fauny, O., and Rahmania, R.: Tracking the Stranded Area of Marine Debris in Indonesian coasts by using Floating Drifter, IOP Conf. Ser.- Earth Environ. Sci., 925, 12034, https://doi.org/10.1088/1755-1315/925/1/012034, 2021.

Robbe, E., Woelfel, J., Balc˘iu¯nas, A., and Schernewski, G.: An Impact Assessment of Beach Wrack and Litter on Beach Ecosystem Services to Support Coastal Management at the Baltic Sea, Environ. Manage., 68, 835-859, https://doi.org/10.1007/s00267-021-01533-3, 2021.

Rodríguez, J. G., Garmendia, J. M., Muxika, I., Gómez-Ballesteros, M., Quincoces, I., Díez, I., Arrese, B., Sánchez, F., and Galparsoro, I.: Macrofaunal variability in the continental shelf and canyons in the southeastern Bay of Biscay, Reg. Stud. Mar. Sci., 48, 102012, https://doi.org/10.1016/j.rsma.2021.102012, 2021.

Rodríguez-Díaz, L., Gómez-Gesteira, J. L., Costoya, X., Gómez-Gesteira, M., and Gago, J.: The Bay of Biscay as a trapping zone for exogenous plastics of different sizes, J. Sea Res., 163, 101929, https://doi.org/10.1016/j.seares.2020.101929, 2020.

Rubio, A., Reverdin, G., Fontán, A., González, M., and Mader, J.: Mapping near-inertial variability in the SE Bay of Biscay from HF radar data and two offshore moored buoys, Geophys. Res. Lett., 38, 1-6, https://doi.org/10.1029/2011GL048783, 2011.

Rubio, A., Solabarrieta, L., Gonzalez, M., Mader, J., Castanedo, S., Medina, R., Charria, G., and Aranda, J. A.: Surface circulation and Lagrangian transport in the SE Bay of Biscay from HF radar data, Ocean. 2013 MTS/IEEE Bergen Challenges North. Dimens., https://doi.org/10.1109/OCEANS-Bergen.2013.6608039, 2013.

Rubio, A., Mader, J., Corgnati, L., Mantovani, C., Griffa, A., Novellino, A., Quentin, C., Wyatt, L., Schulz-Stellenfleth, J., Horstmann, J., Lorente, P., Zambianchi, E., Hartnett, M., Fernandes, C., Zervakis, V., Gorringe, P., Melet, A., and Puillat, I.: HF Radar Activity in European Coastal Seas: Next Steps toward a Pan-European HF Radar Network, Front. Mar. Sci., 4, 8, https://doi.org/10.3389/fmars.2017.00008, 2017.

Rubio, A., Caballero, A., Orfila, A., Hernández-Carrasco, I., Ferrer, L., González, M., Solabarrieta, L., and Mader, J.: Eddy-induced cross-shelf export of high Chl-*a* coastal waters in the SE Bay of Biscay, Remote Sens. Environ., 205, 290-304, https://doi.org/10.1016/j.rse.2017.10.037, 2018.

Rubio, A., Hernández-Carrasco, I., Orfila, A., González, M., Reyes, E., Corgnati, L., Berta, M., Griffa, A., and Mader, J.: A Lagrangian approach to monitor local particle retention conditions in coastal areas, J. Oper. Oceanogr., 13, S1-S172, https://doi.org/10.1080/1755876X.2020.1785097, 2020.

Ruiz, I., Basurko, O. C., Rubio, A., Delpey, M., Granado, I., Declerck, A., Mader, J., and Cózar, A.: Litter Windrows in the South-East Coast of the Bay of Biscay: An Ocean Process Enabling Effective Active Fishing for Litter, Front. Mar. Sci., 7, 308, https://doi.org/10.3389/fmars.2020.00308, 2020.

Ruiz, I., Abascal, A. J., Basurko, O. C., and Rubio, A.: Modelling the distribution of fishing-related floating marine litter within the

Bay of Biscay and its marine protected areas, Environ. Pollut., 292, 118216, https://doi.org/10.1016/j.envpol.2021.118216, 2022a.

Ruiz, I., Rubio, A., Abascal, A. J., and Basurko, O. C.: Supplemental videos of the paper "Modelling floating riverine litter in the south-eastern Bay of Biscay: a regional distribution from a seasonal perspective", TIB [video], https://doi.org/10.5446/s_1355, 2022b.

Russell, K.: Spain's Coastal Authority Uses Spot Trace for Search and Rescue Training, https://www.satellitetoday.com/telecom/2017/06/20/spains-coastal-authority-uses-spot-trace-search-rescue-training/ (last access: 21 December 2019), 2017.

Ryan, P. G.: Does size and buoyancy affect the long-distance transport of floatingdebris?, Environ. Res. Lett., 10, 1, https://doi.org/10.1088/1748-9326/10/8/084019, 2015.

Schmidt, C., Krauth, T., and Wagner, S.: Export of Plastic Debris by Rivers into the Sea, Environ. Sci. Technol., 51, 12246-12253, https://doi.org/10.1021/acs.est.7b02368, 2017.

Schuyler, Q., Wilcox, C., Lawson, T. J., Ranatunga, R. R. M. K. P., Hu, C.-S., Global Plastics Project Partners, and Hardestey, B. D.: Human Population Density is a Poor Predictor of Debris in the Environment, Front. Environ. Sci., 9, 583454, https://doi.org/10.3389/fenvs.2021.583454, 2021.

Sheppard, C.: World Seas: An Environmental Evaluation: Volume I: Europe, The Americas and West Africa, edited by: Sheppard, C., Academic press, Elsevier, https://doi.org/10.1016/C2015-0-04330-1, 2018.

Solabarrieta, L., Rubio, A., Castanedo, S., Medina, R., Charria, G., and Hernández, C.: Surface water circulation patterns in the southeastern Bay of Biscay: New evidences from HF radar data, Cont. Shelf Res., 74, 60-76, https://doi.org/10.1016/j.csr.2013.11.022, 2014.

Solabarrieta, L., Rubio, A., Cárdenas, M., Castanedo, S., Esnaola, G., Méndez, F. J., Medina, R., and Ferrer, L.: Probabilistic relationships between wind and surface water circulation patterns in the SE Bay of Biscay, Ocean Dynam., 65, 1289-1303, https://doi.org/10.1007/s10236-015-0871-5, 2015.

Solabarrieta, L., Frolov, S., Cook, M., Paduan, J., Rubio, A., González, M., Mader, J., and Charria, G.: Skill Assessment of HF Radar–Derived Products for Lagrangian Simulations in the Bay of Biscay, J. Atmos. Ocean. Tech., 33, 2585-2597, https://doi.org/10.1175/jtech-d-16-0045.1, 2016.

Solabarrieta, L., Hernández-Carrasco, I., Rubio, A., Campbell, M., Esnaola, G., Mader, J., Jones, B. H., and Orfila, A.: A new Lagrangian-based short-term prediction methodology for high-frequency (HF) radar currents, Ocean Sci., 17, 755-768, https://doi.org/10.5194/os-17-755-2021, 2021.

Stanev, E. V., Badewien, T. H., Freund, H., Grayek, S., Hahner, F., Meyerjürgens, J., Ricker, M., Schöneich-Argent, R. I., Wolff, J. O., and Zielinski, O.: Extreme westward surface drift in the North Sea: Public reports of stranded drifters and Lagrangian tracking, Cont. Shelf Res., 177, 24-32, https://doi.org/10.1016/j.csr.2019.03.003, 2019.

Teles-Machado, A., Peliz, Á., McWilliams, J. C., Dubert, J., and Cann, B. Le: Circulation on the Northwestern Iberian Margin: Swoddies, Prog. Oceanogr., 140, 116-133, https://doi.org/10.1016/J.POCEAN.2015.09.011, 2016.

Tong, X., Jong, M.-C., Zhang, J., You, L., and Gin, K. Y.-H.: Modelling the spatial and seasonal distribution, fate and transport of floating plastics in tropical coastal waters, J. Hazard. Mater., 414, 125502, https://doi.org/10.1016/j.jhazmat.2021.125502, 2021.

Van Der Mheen, M., Pattiaratchi, C., and Van Sebille, E.: Role of Indian Ocean Dynamics on Accumulation of Buoyant Debris, J. Geophys. Res.-Oceans, 124, 2571-2590, https://doi.org/10.1029/2018JC014806, 2019.

van der Wal, M., van der Meulen, M., Tweehuijsen, G., Peterlin, M., Palatinus, A., Kovacˇ Viršek, M., Coscia, L., and Kržan, A.: Identification and Assessment of Riverine Input of (Marine) Litter, Final Rep. Eur. Comm. DG Environ. under Framew. Contract No. ENV.D.2/FRA/2012/0025, 1-208, 2015.

van Emmerik, T. and Schwarz, A.: Plastic debris in rivers, WIREs Water, 7, e1398, https://doi.org/10.1002/wat2.1398, 2020.

van Sebille, E., England, M. H., and Froyland, G.: Origin, dynamics and evolution of ocean garbage patches from observed surface drifters, Environ. Res. Lett., 7, 044040, https://doi.org/10.1088/1748-9326/7/4/044040, 2012.

Van Sebille, E., Aliani, S., Law, K. L., Maximenko, N., Alsina, J. M., Bagaev, A., Bergmann, M., Chapron, B., Chubarenko, I., and Cózar, A.: The physical oceanography of the transport of floating marine debris, Environ. Res. Lett., 15, 23003, https://doi.org/10.1088/1748-9326/ab6d7d, 2020.

van Sebille, E., Zettler, E., Wienders, N., Amaral-Zettler, L., Elipot, S., and Lumpkin, R.: Dispersion of Surface Drifters in the Tropical Atlantic, Front. Mar. Sci., 7, 607426, https://doi.org/10.3389/fmars.2020.607426, 2021.

van Utenhove, E.: Modelling the transport and fate of buoyant macroplastics in coastal waters, Master thesis, http://resolver.tudelft.nl/uuid:be6a41d2-6071-47b9-926d-f22c23edadba (last access: 25 November 2022), 2019.

Weideman, E. A., Perold, V., Omardien, A., Smyth, L. K., and Ryan, P. G.: Quantifying temporal trends in anthropogenic litter in a rocky intertidal habitat, Mar. Pollut. Bull., 160, 111543, https://doi.org/10.1016/J.MARPOLBUL.2020.111543, 2020.

Wendt-Potthoff, K., Avellán, T., van Emmerik, T., Hamester, M., Kirschke, S., Kitover, D., and Schmidt, C.: Monitoring Plastics in Rivers and Lakes: Guidelines for the Harmonization of Methodologies, edited by: Smith, J., United Nations Environment Programme, ISBN 9789280738193, 2020.

Widyatmoko, A. C., Hardesty, B. D., and Wilcox, C.: Detecting anchored fish aggregating devices (AFADs) and estimating use patterns from vessel tracking data in small-scale fisheries, Sci. Rep.-UK, 11, 17909, https://doi.org/10.1038/s41598-021-97227-1, 2021.

Woods, J. S., Verones, F., Jolliet, O., Vázquez-Rowe, I., and Boulay, A.-M.: A framework for the assessment of marine litter impacts in life cycle impact assessment, Ecol. Indic., 129, 107918, https://doi.org/10.1016/j.ecolind.2021.107918, 2021.

Yoon, J.-H., Kawano, S., and Igawa, S.: Modeling of marine litter drift and beaching in the Japan Sea, Mar. Pollut. Bull., 60, 448, https://doi.org/10.1016/j.marpolbul.2009.09.033, 2010.

Arturo Elosegi editorearen argazkia

Munduko basoen % 40k soilik mantentzen du integritate ekologikoa

Giza jarduerek basoetan eragin duten kaltea kuantifikatzean, munduko bazter gehienetan basoen osasuna kolokan dagoela ikusi da. Basoek berezkoa duten egitura ekologikoa galtzean, planetari eta gizarteari eskaintzen zizkioten onurak ere galtzen dira. Beraz, lehentasunezkoa da integritatea mantentzen duten basoak babesteko neurriak hartzea.

Basoak funtsezkoak dira klima erregulatzeko, biodibertsitatea ahalbidetzeko eta gure ongizate orokorrerako, baina ebaluazio global honek datu kezkagarriak erakutsi ditu: giza jarduerak munduan geratzen diren basoen % 60 eraldatu ditu jada. Giza jardueren ondorioz basoek pairatutako zatiketak eta konektibitate-galerak, besteak beste, arriskuan jartzen dute basoetako ekosistemen integritatea, eta, ondorioz, baita basoek planetari eskaintzen dizkioten onurak ere.

Orain arte deforestazioa izan da kezka-iturri nagusia, baina ikertzaileek argi dute basoek pairatu duten eraldaketa-maila ere erabakigarria dela haien osasunerako. Hainbat giza jarduerak, hala nola urbanizazioak, bestelako giza azpiegiturek, nekazaritzak eta zuhaitz-mozketek, basoek berezkoa duten egitura galtzea ekartzen dute: espezieen osaera, zuhaitzen adin-aniztasuna, lurraren kalitatea eta basoen integritatea osatzen duten bestelako faktoreak. Ondorioz, planetarentzat funtsezkoak diren onurak emateko gaitasuna ere galtzen dute: karbonoa biltegiratzea, planeta ur garbiz hornitzea eta biodibertsitaterako beharrezkoak diren habitatak sortzea.

Europako basoekiko kezka

Ikertzaileen ustez, kezkagarria da Lurreko basoen % 40k baino ez izatea integritate-maila altua. Gainera, baso horiek planetako eskualde jakin batzuetan baino ez daude: batez ere, Kanadako iparraldean, Errusiako iparraldean, Amazonian, Erdialdeko Afrikan eta Ginea Berrian. Gainerako eskualdeek –Europak, Asiako hego-ekialdeak eta Hego Amerikak, besteak beste– integritate txikiko basoak dituzte gehientsuenak.

Ikerketaren emaitzek erakusten dute lehentasunezkoa dela basoen integritatea bermatzeko neurriak hartzea. Hasteko, funtsezkoa da datu esanguratsu batzuk kontuan hartzea: une honetan, integritate handiko basoen % 27 baino ez dago eremu babestuen barruan. Beraz, gainerakoak babesteko neurriak hartu beharko lirateke. Eta, bestetik, eremu babestuetan dauden basoek ere integritate ertaina edo txikia dute, kasu askotan. Gaur egun eremu babestuetan egoteak ez duenez bermatzen basoen integritatea, babes-neurrien egokitasuna berrikusi beharko dela uste dute.

«Integritate handiko basoak kartografiatzeak baso horiek babes-plan berezietan sartzeko balio dezake»

Ikerketak proposatzen du integritate handiko basoak babestea eta integritate ertain/baxukoak lehengoratzea. Horretarako, funtsezkoa da kudeaketarako estrategia eta politika eraginkorrak izatea, eskualde bakoitzaren testuingurura egokituak. Hasteko, integritate handiko bertako basoak kartografiatzea proposatzen dute, gerora baso horien balioa aitortu eta babes-plan berezietan sartzen direla ziurtatzeko. Bestetik, ezinbestekotzat jotzen dute politikariek basoen integritatea hobetzeko helburu zehatzak eta kuantifikagarriak ezartzea, eta gogorarazi dute integritate-galera saihestea gero basoak lehengoratzea baino estrategia hobea dela beti, zaharberritzea garestiagoa eta zailagoa delako.

Ezinbestekoa ikusten dute bide horri ekitea, betiere klima, biodibertsitatea, garapen jasangarria eta justizia soziala helburu dituen Nazio Batuen Erakundearen estrategia globalaren barnean. Bestelakoan, ikertzaileek uste dute ezingo direla bermatu etorkizuneko belaunaldientzat funtsezkoak diren zerbitzuak.

InpAkta 00 (2023) 42-58
https://doi.org/10.1038/s41467-020-19493-3

Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity

H. S. Grantham[1*], A. Duncan[1], T. D. Evans[1], K. R. Jones[1], H. L. Beyer[2], R. Schuster[3], J. Walston[1], J. C. Ray[4], J. G. Robinson[1], M. Callow[1], T. Clements[1], H. M. Costa[1], A. Degemmis[1], P. R. Elsen[1], J. Ervin[5], P. Franco[1], E. Goldman[6], S. Goetz[7], A. Hansen[8], E. Hofsvang[9], P. Jantz[7], S. Jupiter[1], A. Kang[1], P. Langhammer[10, 11], W. F. Laurance[12], S. Lieberman[1], M. Linkie[1], Y. Malhi[13], S. Maxwell[2], M. Mendez[1], R. Mittermeier[10], N. J. Murray[12, 14], H. Possingham[15, 16], J. Radachowsky[1], S. Saatchi[17], C. Samper[1], J. Silverman[1], A. Shapiro[18], B. Strassburg[19], T. Stevens[1], E. Stokes[1], R. Taylor[6], T. Tear[1], R. Tizard[1], O. Venter[20], P. Visconti[21], S. Wang[1] & J. E. M. Watson[1, 2]

[1] Wildlife Conservation Society, Global Conservation Program; [2] School of Earth and Environmental Sciences, University of Queensland; [3] Department of Biology, Carleton University; [4] Wildlife Conservation Society Canada; [5] United Nations Development Programme, One United Nations Plaza; [6] World Resources Institute; [7] Global Earth Observation & Dynamics of Ecosystems Lab, School of Informatics, Computing, and Cyber Systems, Northern Arizona University; [8] Landscape Biodiversity Lab, Ecology Department, Montana State University; [9] Rainforest Foundation Norway; [10] Global Wildlife Conservation; [11] School of Life Sciences, Arizona State University; [12] Centre for Tropical Environmental and Sustainability Science, College of Science and Engineering, James Cook University; [13] Environmental Change Institute, School of Geography and the Environment, University of Oxford; [14] College of Science and Engineering, James Cook University; [15] School of Biological Sciences, University of Queensland; [16] The Nature Conservancy; [17] Jet Propulsion Laboratory, California Institute of Technology; [18] World Wide Fund for Nature Germany, Space+Science; [19] International Institute of Sustainability; [20] Natural Resource and Environmental Studies Institute, University of Northern British Columbia; [21] International Institute for Applied Systems Analysis.

ABSTRACT: Many global environmental agendas, including halting biodiversity loss, reversing land degradation, and limiting climate change, depend upon retaining forests with high ecological integrity, yet the scale and degree of forest modification remain poorly quantified and mapped. By integrating data on observed and inferred human pressures and an index of lost connectivity, we generate a globally consistent, continuous index of forest condition as determined by the degree of anthropogenic modification. Globally, only 17.4 million km^2 of forest (40.5%) has high landscape-level integrity (mostly found in Canada, Russia, the Amazon, Central Africa, and New Guinea) and only 27% of this area is found in nationally designated protected areas. Of the forest inside protected areas, only 56% has high landscape-level integrity. Ambitious policies that prioritize the retention of forest integrity, especially in the most intact areas, are now urgently needed alongside current efforts aimed at halting deforestation and restoring the integrity of forests globally.

* **Corresponding author / Harremanetan jartzeko:** H. S. Grantham. Wildlife Conservation Society, Global Conservation Program, Bronx, New York 10460, USA. e-mail: hgrantham@wcs.org. **Contributed by** Arturo Elosegi **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Grantham, H. S. *et al.* «Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity», *Nature Communications* | (2020)11:5978 (https://doi.org/10.1038/s41467-020-19493-3)

Deforestation is a major environmental issue [1], but far less attention has been given to the degree of anthropogenic modification of remaining forests, which reduces ecosystem integrity and diminishes many of the benefits that these forests provide [2, 3]. This is worrying since modification is potentially as significant as outright forest loss in determining overall environmental outcomes [4]. There is increasing recognition of this issue, for forests and other ecosystems, in synthesis reports by global science bodies such as the global assessment undertaken by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [5], and it is now essential that the scientific community develop improved tools and data to facilitate the consideration of levels of integrity in decision-making. Mapping and monitoring this globally will provide essential information for coordinated global, national, and local policy-making, planning, and action, to help nations and other stakeholders achieve the Sustainable Development Goals (SDGs) and implement other shared commitments such as the United Nations Convention on Biological Diversity (CBD), Convention to Combat Desertification (UNCCD), and Framework Convention on Climate Change (UNFCCC).

Ecosystem integrity is foundational to all three of the Rio Conventions (UNFCCC, UNCCD, CBD) [6]. As defined by Parrish *et al.* [7], it is essentially the degree to which a system is free from anthropogenic modification of its structure, composition, and function. Such modification causes the reduction of many ecosystem benefits, and is often also a precursor to outright deforestation [8, 9]. Forests largely free of significant modification (i.e., forests having high ecosystem integrity), typically provide higher levels of many forest benefits than modified forests of the same type [10], including; carbon sequestration and storage [11], healthy watersheds [12], traditional forest use [13], contribution to local and regional climate processes [14], and forest-dependent biodiversity [15-18]. Industrial-scale logging, fragmentation by infrastructure, farming (including cropping and ranching) and urbanization, as well as less visible forms of modification such as over-hunting, wood fuel extraction, and changed fire or hydrological regimes [19, 20], all degrade the degree to which forests still support these benefits, as well as their long-term resilience to climate change [10]. There can be trade-offs, however, between the benefits best provided by less-modified forests (e.g., regulatory functions such as carbon sequestration) and those production services that require some modification (e.g., timber production). These trade-offs can, at times, result in disagreement among stakeholders as to which forest benefits should be prioritized [21].

In recent years, easily accessible satellite imagery and new analytical approaches have improved our ability to map and monitor forest extent globally [22-24]. However, while progress has been made in developing tools for assessment of global forest losses and gains, consistent monitoring of the degree of forest modification has proved elusive [25, 26].

Technical challenges include the detection of low intensity and unevenly distributed forest modification, the wide diversity of changes that comprise forest modification, and the fact that many changes are concealed by the forest canopy [25]. New approaches are emerging on relevant forest indicators, such as canopy height, canopy cover and fragmentation, and maps of different human pressures, which are used as proxies for impacts on forests [27-30]. Some binary measures of forest modification, such as Intact Forest Landscapes [31] and wilderness areas [32], have also been mapped at the global scale and used to inform policy, but do not resolve the degree of modification within remaining forests, which we aimed to do with this assessment.

Human activities influence the integrity of forests at multiple spatial scales, including intense, localized modifications such as road-building and canopy loss, more diffuse forms of change that are often spatially associated with these localized pressures (e.g., increased accessibility for hunting, other exploitation, and selective logging), and changes in spatial configuration that alter landscape- level connectivity. Previ-

ous studies have quantified several of these aspects individually [27-29], but there is a need to integrate them to measure and map the overall degree of modification considering these landscape-level anthropogenic influences at each site. Here, we integrate data on forest extent defined as all woody vegetation taller than 5 m, following [23], observed human pressures (e.g., infrastructure) which can be directly mapped using current datasets, other inferred human pressures (e.g., collection of forest materials) that occur in association with those that are observed but cannot be mapped directly, and alterations in forest connectivity, to create the Forest Landscape Integrity Index (FLII), that describes the degree of forest modification for the beginning of 2019 (Fig. 1).

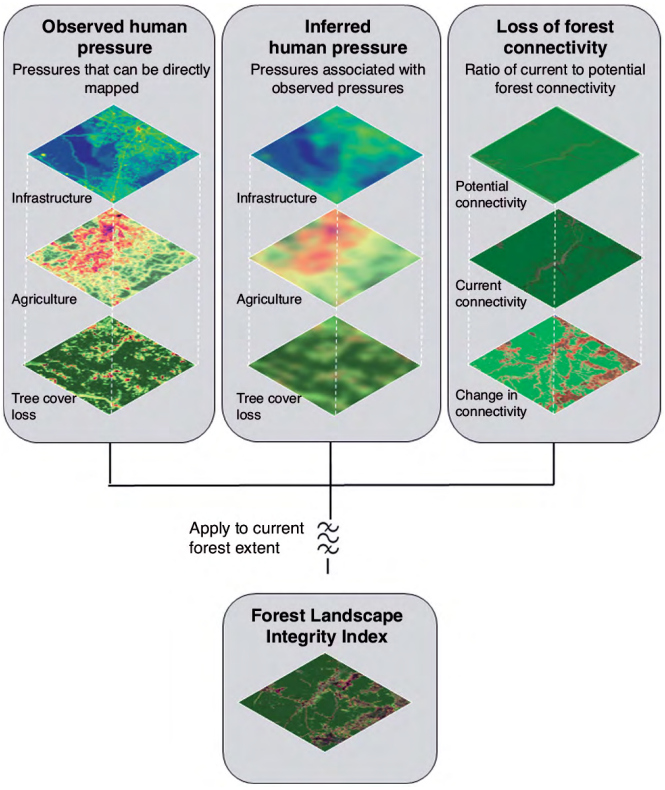


Figure 1. Methods used to construct the Forest Landscape Integrity Index. The Forest Landscape Integrity Index was constructed based on three main data inputs: (1) observed pressures (infrastructure, agriculture, tree cover loss), (2) inferred pressure modeled based on proximity to the observed pressures, and (3) change in forest connectivity

The result is a globally applicable, continuous-measure map of landscape-level forest integrity (hereafter, integrity), which offers a timely indicator of the status and management needs of Earth's remaining forests. The results show there has been a huge loss of forest integrity. To give a global overview we summarize the results according to three simple, illustrative categories of integrity (which we term "high", "medium", and "low") while noting that the underlying continuous index enables much finer distinctions to be made for detailed analysis in diverse contexts. This reveals around 40% of remaining forests have high forest integrity. Further, our methodological framework (Fig. 1) can be adapted to match local conditions at national or subnational scales and for different weightings to be applied.

1. Results

Forest modification caused by human activity is both highly pervasive and highly variable across the globe (Fig. 2). We found 31.2% of forests worldwide are experiencing some form of observed human pressure, which included infrastructure, agriculture, and recent deforestation. Our models also inferred the likely occurrence of other pressures, and the impacts of lost connectivity, in almost every forest location (91.2% of forests), albeit sometimes at very low levels. Diverse, recognizable patterns of forest integrity can be observed in our maps at a range of scales, depending on the principal forms and general intensity of human activity in an area. Broad regional trends can be readily observed, for example, the overall gradient of decreasing human impact moving northwards through eastern North America (Fig. 2), and finer patterns of impact are also clearly evident, down to the scale of individual protected areas, forest concessions, settlements, and roads (Supplementary Fig. 2).

FLII scores range from 0 (lowest integrity) to 10 (highest). We discretized this range to define three broad illustrative categories: low (≤6.0); medium (>6.0 and <9.6); and high integrity (≥9.6) by

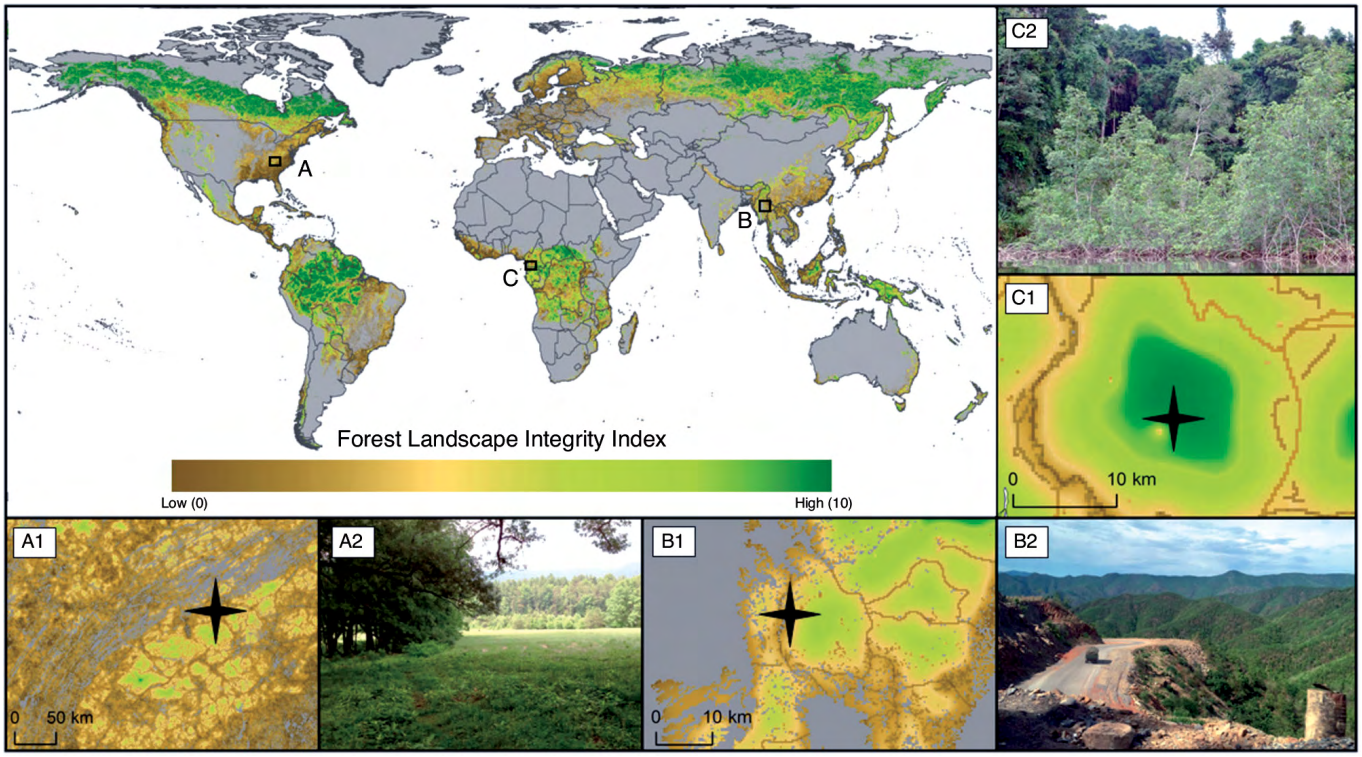


Figure 2. Forest Landscape Integrity Index map. A global map of Forest Landscape Integrity for the start of 2019. Three regions are highlighted including (a) Smoky Mountains National Park in Tennessee USA, (b) a region in Shan State Myanmar, and (c) Reserva Natural del Estuario del Muni in Equatorial Guinea. Maps A1-C1 shows the Forest Landscape Integrity Index for these locations. A2, B2, and C2 are photographs from within these regions: (A2) the edge of Smoky Mountains National Park; (B2) shows a logging truck passing through some partially degraded forest along a newly constructed highway in Shan Stat; and, (C3) shows an intact mangrove forest within Reserva Natural del Estuario del Muni, near the border with Gabon. The stars in (a), (b), and (c) indicate approximate location of where these photos were taken. All photos were taken by H.S.G.

benchmarking against reference locations worldwide (see Methods, Supplementary Table 4). Only 40.5% (17.4 million km^2) of the forest was classified as having high integrity (fig. 3; Table 1). Moreover, even in this category of high integrity 36% still showed at least a small degree of human modification. The remaining 59% (25.6 million km^2) of the forest was classified as having low or medium integrity, including 25.6% (11 million km^2) with low integrity (Fig. 3; Table 1). When we analyzed across biogeographical realms defined by [33] not a single biogeographical realm of the world had more than half of its forests in the high category (Fig. 3; Table 1).

The biogeographical realms with the largest area of forest with high integrity are the Paleartic, particularly northern Russia, and the Neartic, in northern Canada, Rocky Mountains, and Alaska (Fig. 3). There are also large areas of forest with high integrity in the Neotropics, concentrated in the Amazon region, including within the Guianas, Atlantic forest in Brazil, southern Chile, and parts of Mesoamerica (Fig. 3, Table 1). The Afrotropic realm has significant areas with high integrity, particularly within the humid forests of central Africa (e.g., in Republic of Congo and Gabon) and in some of the surrounding drier forest/woodland belts (e.g., in South Sudan, Angola, and Mozambique) (Fig. 3). Some smaller patches occur in West Africa and Madagascar. In tropical Asia-Pacific, the largest tracts of forest with high integrity are in New Guinea. Smaller but still very significant tracts of forest with high integrity are also scattered elsewhere in each of the main forested regions, including parts of Sumatra, Borneo, Myanmar, and other parts of the Greater Mekong subregion.

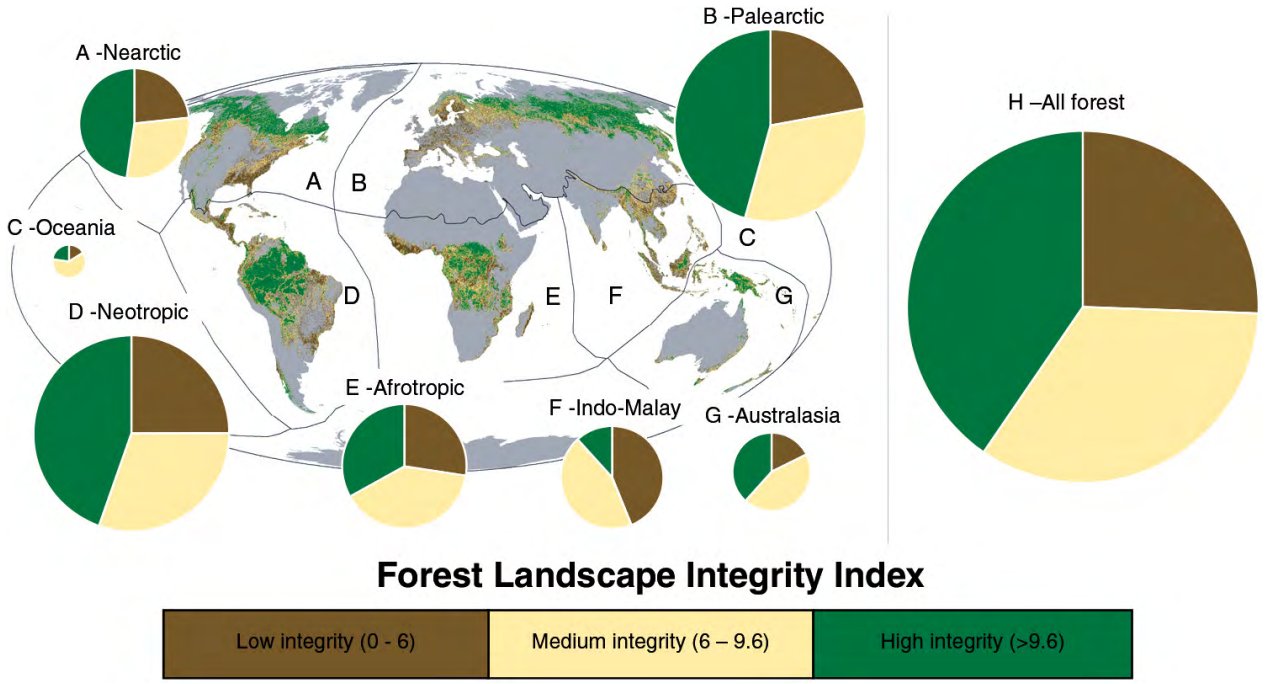


Figure 3. Forest Landscape Integrity Index map categorized into three illustrative classes. The Forest Landscape Integrity Index for 2019 categorized into three broad, illustrative classes and mapped across each biogeographic realm (a-g). The size of the pie charts indicates the relative size of the forests within each realm (a-g), and h shows all the world's forest combined

Concentrations of the forest with low integrity are found in many regions including west and central Europe, the south-eastern USA, island and mainland South-East Asia west of New Guinea, the Andes, much of China and India, the Albertine Rift, West Africa, Mesoamerica, and the Atlantic Forests of Brazil (Fig. 3). The overall extent of forests with low integrity is greatest in the Palearctic realm, followed by the Neotropics, which are also those biogeographic realms with the largest forest cover (Table 1). The Indo-Malayan realm has the highest percentage with low integrity, followed by the Afrotropics (Fig. 3; Table 1).

Table 1
Brief title: Forest Landscape Integrity Index scores for each biogeographic realm

Biogeographic realm	Historical forest area (km²)	Current forest area (km²)	Proportion of forest remaining (%)	FLII (mean)	High (9.6-10)		Medium (6-9.6)		Low (0-6)	
					km²	% of realm	km²	% of realm	km²	% of realm
Afrotropic	9,071,897	7,362,740	81.2	7.34	2,450,953	33.3	2,903,483	39.4	2,008,304	27.3
Australasia	2,225,054	1,711,684	76.9	8.05	656,701	38.4	753,188	44,0	301,796	17.6
Indo-malayan	4,797,518	3,596,249	75.0	5.9	420,977	11.7	1,599,049	44.5	1,576,223	43.8
Neotropic	14,965,342	10,271,519	68.6	7.81	4,579,406	44.6	3,122,706	30.4	2,569,407	25,0
Oceania	30,746	23,389	76.1	7.66	5,279	22.6	14,331	61.3	3,780	16.2
Palearctic	16,524,088	12,172,668	73.7	8,00	5,571,997	45.8	3,910,629	32.1	2,690,042	22.1
Nearctic	9,756,589	7,794,117	79.9	7.84	3,716,855	47.7	2,257,518	29,0	1,819,744	23.3
Total	57,371,234	42,932,367	74.8	7.76	17,402,170		14,560,903		10,969,294	

A summary of the Forest Landscape Integrity Index scores for each biogeographic realm globally, measuring the mean score, in addition to the area and proportion of realm for each category of integrity. Scores are divided into three categories of integrity: high, medium, and low.

These patterns result in variation of forest integrity scores in ways that allow objective comparisons to be made between locations and at a resolution relevant for policy and management planning, such as at national and sub-national scales. The global average FLII score is 7.76 (Table 1), representing a medium level of integrity. However, the average score across countries, disregarding their size, is 5.48, suggesting that low scores dominate in many of the smaller countries, and indeed a quarter of forested countries have a national average score < 4. National mean scores vary widely, ranging from >9 in Guyana, French Guiana, Gabon, Sudan, and South Sudan to <3 in Sierra Leone and many west European countries (see Fig. 4. and Supplementary Table 5 for a full list of countries). Provinces and other sub-national units vary even more widely (see Supplementary Fig. 2 and Supplementary Table 6).

Over one-quarter (26.1%) of all forests with high integrity fall within protected areas, compared to just 13.1% of low and 18.5% of medium integrity forests respectively. For all forests that are found within nationally designated protected areas (around 20% of all forests globally), we found the proportions of low, medium, and high integrity forests were 16.8%, 30.3%, and 52.8%, respectively (Table 2). Within the different protected area categories, we typically found that there was more area within the high integrity category versus the medium and low except for Category V (protected landscape/seascape) (Table 2). However, with 47.1% of forests within protected areas having low to medium integrity overall, it is clear that forests considered protected are already often fairly modified (Table 2). Even though they are quite modified, some of these forests might still have high conservation importance, such as containing endangered species.

Table 2
Brief title: Forest Landscape Integrity Index scores for different types of protected areas

Protected area category	Total forest (km^2)	FLII (mean)	High (score 9.6-10)		Medium (score 6-9.6)		Low (score 0-6)	
			km^2	% of protected area	km^2	% of protected area	km^2	% of protected area
Ia (strict nature reserve)	439,082	9.27	304,329	69.31	106,703	24.30	28,049	6.39
Ib (wilderness area)	367,330	9.22	240,453	65.46	102,096	27.79	24,780	6.75
II (national park)	1,900,000	9.14	1,223,138	64.38	540,805	28.46	136,056	7.16
III (natural monument or feature)	113,805	8.49	54,476	47.87	40,021	35.17	19,308	16.97
IV (habitat/species management area)	838,707	8.69	432,828	51.61	268,027	31.96	137,850	16.44
V (protected landscape/seascape)	840,919	6.40	224,491	26.70	295,769	35.17	320,658	38.13
VI (Protected area with sustainable use of natural resources)	1,472,278	9.21	1,026,169	69.70	344,617	23.41	101,491	6.89
Not Applicable / Not Assigned / Not Reported	2,613,541	8.29	1,030,430	39.42	906,745	34.69	676,365	25.88
All Protected Areas	8,585,661	8.55	4,536,314	52.83	2,694,784	30.34	1,444,562	16.82

A summary of the Forest Landscape Integrity Index scores for each type of protected area designation based on the IUCN Protected Areas categories measuring mean score, in addition to the area and proportion of realm for each category of integrity. Scores are divided into three categories of integrity: high, medium, and low.

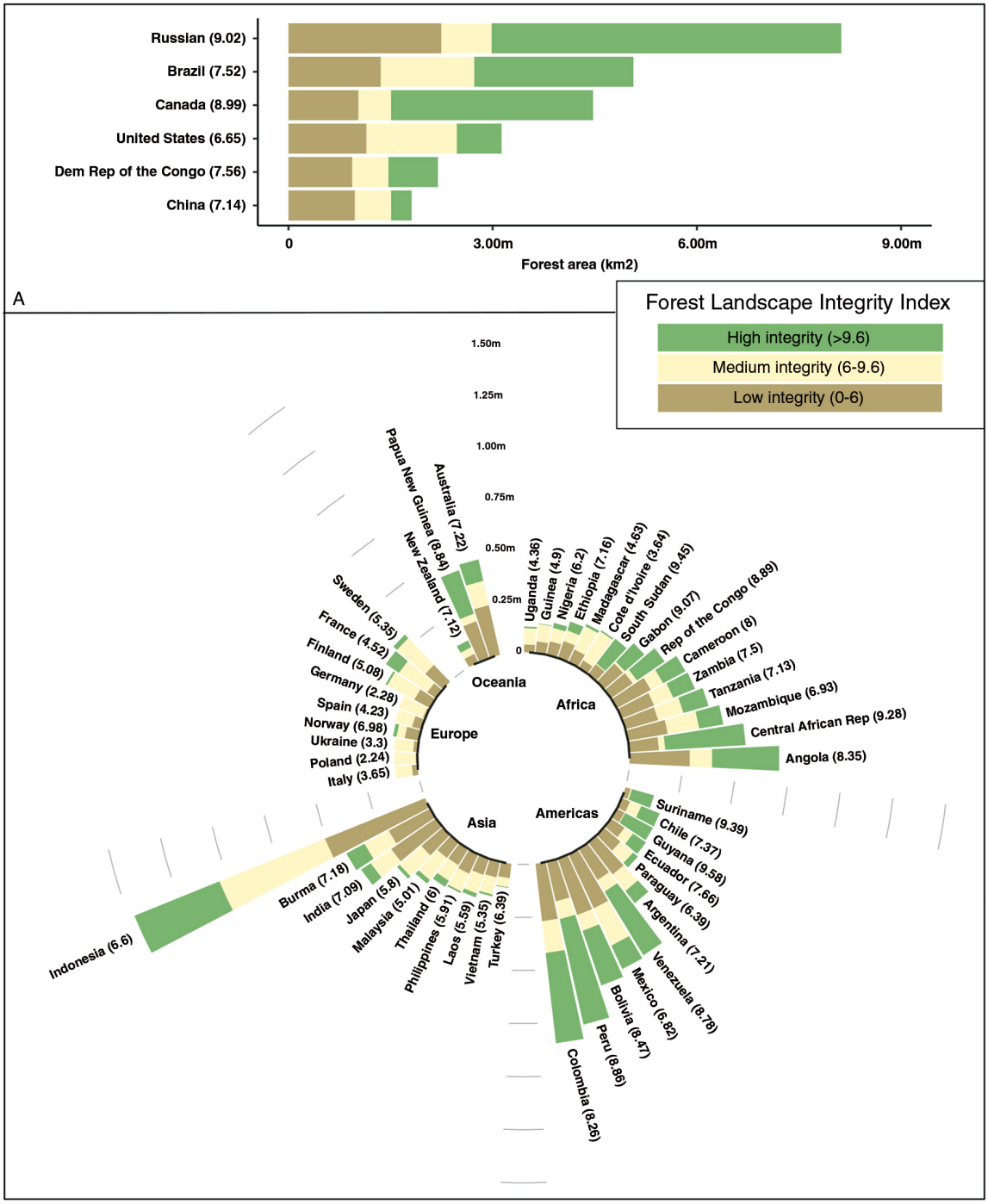


Figure 4. Forest Landscape Integrity Index map categorized into three illustrative classes for each major forested country. The Forest Landscape Integrity Index for 2019 categorized into three broad, illustrative classes for each major forested country in the world. (a) countries with a forest extent larger than 1 million km², and (b) countries with forest extent between 1 million km² and 100,000 km² of forest. The size of the bar represents the area of a country's forests

2. Discussion

By providing a transparent and defensible methodological framework, and by taking advantage of global data on forest extent, human drivers of forest modification, and changes in forest connectivity, our analysis paints a sobering picture of the extent of human impacts on the world's forests. This analysis enables the changes that degrade many forest values to be visualized in a way for policymakers and deci-

sion-makers to see where forests that survive in good condition are found. By integrating data on multiple human pressures that are known to modify forests, our analysis moves global quantification beyond the use of simple categories, or solely using pressure indicators as proxies for integrity, to a more nuanced depiction of this issue as a continuum, recognizing that not all existing forests are in the same condition. Our analysis reveals that severe and extensive forest modification has occurred across all biogeographic regions of the world. Consequently, indices only using forest extent may inadequately capture the true impact of human activities on forests, and are insensitive to many drivers of forest modification and the resulting losses of forest benefits.

A plan is clearly needed to put in place retention strategies for the remaining forests with high integrity, tailored towards the context in each country or jurisdiction and its different forest types [34-36], because such areas are known to hold exceptional value. Avoiding the loss of integrity is a better strategy than aiming to restore forest condition after it is lost, because restoration is more costly, has a risk of failure, and is unlikely to lead to full recovery of benefits [5]. For the forests with the highest integrity to be retained they should ideally be mapped using nationally appropriate criteria by the countries that hold them, formally recognized, prioritized in spatial plans, and placed under effective management (e.g., protected areas and other effective conservation areas, lands under Indigenous control, etc.). These forests must be protected from industrial development impacts that degrade them through sensible public and private sector policy that is effective at relevant scales [13, 37]. Our global assessment reveals where these places are found, and can be refined at more local scales where better data are available.

Around a third of global forests had already been cleared by 2000 [38], and we show that at least 59% of what remains has low or medium integrity, with > 50% falling in these two broad categories in every biogeographical realm. These levels of human modification result partly from the large areas affected by relatively diffuse anthropogenic pressures whose presence is inferred near forest edges, and by lost connectivity. We also map a surprising level of more localized, observed pressures, such as infrastructure and recent forest loss, which are seen in nearly a third of forested pixels worldwide.

Conservation strategies in these more heavily human-modified forests should focus on securing any remaining fragments of forests in good condition, proactively protecting those forests most vulnerable to further modification [8] and planning where restoration efforts might be most effective [39-41]. In addition, effective management of production forests is needed to sustain yields without further worsening their ecological integrity [42]. More research is required on how to prioritize, manage, and restore forests with low to medium integrity [41, 43], and the FLII presented here might prove useful for this, for example, by helping prioritize where the best returns on investment are, in combination with other sources of data (e.g., carbon) [44].

Loss of forest integrity severely compromises many benefits of forests that are central to achieving multiple Sustainable Development Goals and other societal targets [45, 46]. Therefore, governments must adopt policies and strategies to retain and restore the ecological integrity of their forests, whilst ensuring that the solutions are also economically viable, socially equitable, and politically acceptable within complex and highly diverse local contexts. This is an enormous challenge and our efforts to map the degree of forest modification are designed both to raise awareness of the importance of the issue, and to support implementation through target setting, evidence-based planning, and enhanced monitoring efforts.

Whilst policy targets for halting deforestation are generally precise and ambitious, only vague targets are typically stipulated around reducing levels of forest modification [10, 47]. We urgently need SMART (specific, measurable, achievable, realistic, and time-bound) goals, targets, and indicators for maintaining and restoring forest integrity that directly feeds into higher-level biodiversity, climate, land degradation, and sustainable development goals [48]. Forest specific targets could be included within an over-arching target on ecosystems within the post-2020 Global Biodiversity Framework, which is currently being negotiated among Parties

to the CBD [49]. This target needs to be outcome-focused and address both the extent and the integrity of ecosystems (e.g., using FLII for forests), in a way that enables quantitative, measurable goals to be set and reported on, but allows flexibility for implementation between Parties. The index we provide here could be easily updated annually and utilized by nations as a way to report the state of their forests.

In addition to broader goals in global frameworks, the retention and restoration of forest integrity should also be addressed in nationally-defined goals embodied in, and aligned between, Nationally Determined Contributions under the UNFCCC, efforts to stop land degradation and achieve land degradation neutrality under the UNCCD, and National Biodiversity Strategy and Action Plans under the CBD. Since no single metric can capture all aspects of a country's environmental values, efforts to conserve high levels of forest integrity should be complemented by consideration of areas that support important values according to other measures (e.g., Key Biodiversity Areas [50] and notable socio-cultural landscapes).

A key management tool for maintaining and improving forest integrity is protected areas [10]. We found over a quarter of forests with high integrity are within protected areas, showing that this importance has been widely recognized by some national authorities. However, we also found that nearly half of the forests within protected areas have medium or low integrity. This result aligns with other studies such as Jones *et al.* [51] that found a third of protected areas had high human pressure within them. Compared with more restricted protected areas (e.g., category I), there was a broad trend of decreasing forest integrity in protected area categories that allows more human use, with particularly low mean scores and high percentages of the forest with low integrity in Category V (Protected Landscapes/Seascapes). The exception is category VI, which includes indigenous and community protected areas, some of which contain very extensive areas with low human population pressure, and for which mean integrity scores are comparable to those in category I. Some of these differences probably represent differences at the time of establishment, so time series or quasi-experimental methods are needed to clarify the degree to which the various categories are effective in mitigating threats to integrity, as suggested by Fa *et al.* [52].

The overall level and pervasiveness of impacts on Earth's remaining forests is likely even more severe than our findings suggest, because some input data layers, despite being the most comprehensive available, are still incomplete as there are lags between increases in human pressures and our ability to capture them in spatial datasets e.g., infrastructure [53, 54], (see also Supplementary Note 5 and Supplementary Fig. 1). For example, roads and seismic lines used for natural resource exploration and extraction in northern boreal regions of Canada, are not fully reflected in global geospatial datasets (Supplementary Fig. 1; see also [55]) The over-exploitation of high socio-economic value animals and plants may be quite varied across nations and region, driven by complex social, cultural, economic and governance factors e.g. [56, 57], which are difficult to model spatially but as these data become available, they could be included in further updates of the index. Adding a temporal dimension of the index is an important next step, as it will be possible to start to assess the drivers and underlying caused leading to intact forest erosion which clearly requires further research attention. Furthermore, because natural fires are such an important part of the ecology of many forest systems (e.g., boreal forests) and it is not possible to consistently identify anthropogenic fires from natural fires at a global scales [58] we have taken a strongly conservative approach to fire in our calculations, treating all tree cover loss in 10 km pixels where fire was the dominant driver as temporary, and not treating such canopy loss as evidence of observed human pressure. Varying these assumptions where human activity is shown to be causing permanent tree cover losses, increasing fire return frequencies, or causing fire in previously fire-free systems would result in lower forest extent and/or lower forest integrity scores in some regions than we report.

We map forest integrity based on quantifiable processes over the recent past (since 2000). In some areas modification that occurred prior to this (e.g., historical logging) is not detectable by our methods

but may have influenced the present-day integrity of the forest so, in such cases, we may overestimate forest integrity. This is another reason why our index should be considered as conservative, and we, therefore, recommend that the index be used alongside other lines of evidence to determine the absolute level of the ecological integrity of a given area. Moreover, the definition of forest in this study is all woody vegetation taller than 5 m, following [23] and hence includes not only naturally regenerated forests but also tree crops, planted forests, wooded agroforests, and urban tree cover in some cases. Users should be mindful of this when interpreting the results, especially when observing areas with low forest integrity scores. Inspection of the results for selected countries with reliable plantation maps [59] shows that the great majority of planted forests have low forest integrity scores, because they are invariably associated with dense infrastructure, frequent canopy replacement, and patches of farmland.

We note our measure of forest integrity does not address past, current, and future climate change. As climate change affects forest conditions both directly and indirectly, this is a clear shortfall and needs research attention. The same is true for invasive species, as there are no globally coherent data on the ranges of those invasive species that degrade forest ecosystems, although this issue is indirectly addressed since the presence of many invasive species is likely spatially correlated with the human pressures that we use as drivers in our model [27]. We estimated the likely occurrence of damage caused by inferred pressures using a distance function; this function could be tailored to particular contexts, such as the presence of high-value species or unusually difficult terrain, if training data were available. As global data become available it would also be valuable to incorporate data on other drivers of forest integrity loss. Future research might enable the inclusion of governance effectiveness as a factor in our model, because there are potentially contexts (e.g., well-managed protected areas and community lands, production forests under "sustainable forest management") where the impacts associated with the human pressures we base our map on are at least partially ameliorated [42], and enhanced governance is also likely to be a significant component of some future strategies to maintain and enhance forest integrity.

The framework we present is now being tailored for use at smaller scales, ranging from regional to national and sub-national scales, and even to individual management units, through the development of a cloud-based online tool. Forest definitions and the relative weights of the global parameters we use can be adjusted to fit local contexts and, in many cases, better local data could be substituted, or additional variables incorporated. This would not only increase the precision of the index in representing local realities, but also the degree of ownership amongst national and local policymakers and stakeholders whose decisions are so important in determining forest management trajectories.

3. Methods

To produce our global Forest Landscape Integrity Index (FLII), we combined four sets of spatially explicit datasets representing: (i) forest extent [23]; (ii) observed pressure from high impact, localized human activities for which spatial datasets exist, specifically: infrastructure, agriculture, and recent deforestation [27]; (iii) inferred pressure associated with edge effects [27], and other diffuse processes, (e.g., activities such as hunting and selective logging) [27] modeled using proximity to observed pressures; and iv) anthropogenic changes in forest connectivity due to forest loss [27] (see Supplementary Table 1 for data sources). These datasets were combined to produce an index score for each forest pixel (300 m), with the highest scores reflecting the highest forest integrity (Fig. 1), and applied to forest extent for the start of 2019. We use globally consistent parameters for all elements (i.e., parameters do not vary geographically). All calculations were conducted in Google Earth Engine (GEE) [60].

3.1. Forest extent

We derived a global forest extent map for 2019 by subtracting from the Global Tree Cover product for 2000 [23] annual Tree Cover Loss 2001-2018, except

for losses categorized by Curtis and colleagues [24] as those likely to be temporary in nature (i.e., those due to fire, shifting cultivation and rotational forestry). We applied a canopy threshold of 20% based on related studies e.g. [31, 61], and resampled to 300m resolution and used this resolution as the basis for the rest of the analysis (see Supplementary Note 1 for further methods).

3.2. *Observed human pressures*

We quantify observed human pressures (P) within a pixel as the weighted sum of impact of infrastructure (I; representing the combined effect of 41 types of infrastructure weighted by their estimated general relative impact on forests (Supplementary Table 3), agriculture (A) weighted by crop intensity (indicated by irrigation levels), and recent deforestation over the past 18 years (H; excluding deforestation from fire, see Discussion). Specifically, for pixel *i*:

$$Pi = \exp(-\beta_1 I_i) + \exp(-\beta_2 A_i) + \exp(-\beta_3 H_i) \quad (1)$$

whereby the values of β were selected so that the median of the non-zero values for each component was 0.75. This use of exponents is a way of scaling variables with non-commensurate units so that they can be combined numerically, while also ensuring that the measure of observed pressure is sensitive to change (increase or decrease) in the magnitude of any of the three components, even at large values of I, A, or H. This is an adaptation of the Human Footprint methodology [62]. See Supplementary Note 3 for further details.

3.3. *Inferred human pressures*

Inferred pressures are the diffuse effects of a set of processes for which directly observed datasets do not exist, that include microclimate and species interactions relating to the creation of forest edges [63] and a variety of intermittent or transient anthropogenic pressures such as selective logging, fuelwood collection, hunting; spread of fires and invasive species, pollution, and livestock grazing [64-66]. We modeled the collective, cumulative impacts of these inferred effects through their spatial association with observed human pressure in nearby pixels, including a decline in effect intensity according to distance, and partitioning into stronger short-range and weaker long-range effects. The inferred pressure (P′) on pixel *i* from source pixel *j* is:

$$P'_{i,j} = P_j (w_{i,j} + v_{i,j}) \quad (2)$$

where $w_{i,j}$ is the weighting given to the modification arising from short-range pressure, as a function of distance from the source pixel, and $v_{i,j}$ is the weighting given to the modification arising from long-range pressures.

Short-range effects include most of the processes listed above, which together potentially affect most biophysical features of a forest, and predominate over shorter distances. In our model, they decline exponentially, approach zero at 3 km, and are truncated to zero at 5 km (see Supplementary Note 4).

$$\begin{aligned} w_{i,j} &= \alpha \exp(-\lambda d_{i,j}) && [\text{for } d_{i,j} \le 5 \text{ km}] \\ w_{i,j} &= 0 && [\text{for } d_{i,j} > 5 \text{ km}] \end{aligned} \quad (3)$$

where α is a constant set to ensure that the sum of the weights across all pixels in the range is 1.85 (see below), λ is a decay constant set to a value of 1 (see [67] and other references in Supplementary Note 4) and $d_{i,j}$ is the Euclidean distance between the centers of pixels *i* and *j* expressed in units of km.

Long-range effects include over-exploitation of high socio-economic value animals and plants, changes to migration and ranging patterns, and scattered fire and pollution events. We modeled long-range effects at a uniform level at all distances below 6 km and they then decline linearly with distance, conservatively reaching zero at a radius of 12 km [65, 68] (and other references in Supplementary Note 4):

$$\begin{aligned} V_{i,j} &= \gamma && [\text{for } d_{i,j} \le 6 \text{ km}] \\ v_{i,j} &= \gamma (12 - d_{i,j})/6 && [\text{for } 6 \text{ km} < d_{i,j} \le 12 \text{ km}] \\ v_{i,j} &= 0 && [\text{for } d_{i,j} > 12 \text{ km}] \end{aligned} \quad (4)$$

where γ is a constant set to ensure that the sum of the weights across all pixels in the range is 0.15 and di,j is the Euclidean distance between the centers of pixels *i* and *j*, expressed in kilometers.

The form of the weighting functions for short- and long-range effects and the sum of the weights

($\alpha + \gamma$) were specified based on a hypothetical reference scenario where a straight forest edge is adjacent to a large area with uniform human pressure, and ensuring that in this case total inferred pressure immediately inside the forest edge is equal to the pressure immediately outside, before declining with distance. γ is set to 0.15 to ensure that the long-range effects conservatively contribute no more than 5% to the final index in the same scenario, based on expert opinion and supported e.g., Berzaghi *et al.* [69] regarding the approximate level of impact on values that would be affected by severe defaunation and other long-range effects.

The aggregate effect from inferred pressures (Q) on pixel *i* from all *n* pixels within range (*j* = 1 to *j* = *n*) is then the sum of these individual, normalized, distance-weighted pressures, i.e.,

$$Q_i = \sum_{j=1}^{n} P'_{i,j} \tag{5}$$

3.4. *Loss of forest connectivity*

Average connectivity of forest around a pixel was quantified using a method adapted from Beyer *et al.* [70]. The connectivity C_i around pixel *i* surrounded by *n* other pixels within the maximum radius (numbered *j* = 1, 2**...,** *n*) is given by:

$$C_i = \sum_{j=1}^{n} \left(F_j G_{i,j} \right) \tag{6}$$

where F_j is the forest extent is a binary variable indicating if forested (1) or not (0) and $G_{i,j}$ is the weight assigned to the distance between pixels *i* and *j*. $G_{i,j}$ uses a normalized Gaussian curve, with $\sigma = 20$ km and distribution truncated to zero at 4σ for computational convenience (see Supplementary Note 2). The large value of σ captures landscape connectivity patterns operating at a broader scale than processes captured by other data layers. C_i ranges from 0 to 1 ($C_i \in [0,1]$).

Current Configuration (CC_i) of forest extent in pixel *i* was calculated using the final forest extent map and compared to the Potential Configuration (PC) of forest extent without extensive human modification, so that areas with naturally low connectivity, e.g., coasts and natural vegetation mosaics, are not penalized. PC was calculated from a modified version of the map of Laestadius *et al.* [38] and resampled to 300m resolution (see Supplementary Note 2 for details). Using these two measures, we calculated Lost Forest Configuration (LFC) for every pixel as:

$$LFC_i = 1 - (CC_i / PC_i) \tag{7}$$

Values of $CC_i/PC_i > 1$ are assigned a value of 1 to ensure that LFC is not sensitive to apparent increases in forest connectivity due to inaccuracy in estimated potential forest extent – low values represent least loss, high values greatest loss ($LFC_i \in [0,1]$).

3.5. *Calculating the Forest Landscape Integrity Index*

The three constituent metrics, LFC, P, and Q, all represent increasingly modified conditions the larger their values become. To calculate a forest integrity index in which larger values represent less degraded conditions we, therefore, subtract the sum of those components from a fixed large value (here, 3). Three was selected as our assessment indicates that values of LFC + P + Q of 3 or more correspond to the most severely degraded areas. The metric is also rescaled to a convenient scale (0-10) by multiplying by an arbitrary constant (10/3). The FLII for forest pixel *i* is thus calculated as:

$$FLII_i = [10/3]\,(3 - \min(3, [P_i + Q_i + LFC_i])) \tag{8}$$

where $FLII_i$ ranges from 0 to 10, forest areas with no modification detectable using our methods scoring 10 and those with the most scoring 0.

3.6. *Illustrative forest integrity classes*

Whilst a key strength of the index is its continuous nature, the results can also be categorized for a range of purposes. In this paper three illustrative classes were defined, mapped, and summarized to give an overview of broad patterns of integrity in the world's forests. The three categories were defined as follows.

High Forest Integrity (scores ≥ 9.6) Interiors and natural edges of more or less unmodified naturally regenerated (i.e., non-planted) forest ecosystems,

comprised entirely or almost entirely of native species, occurring over large areas either as continuous blocks or natural mosaics with non-forest vegetation; typically little human use other than low-intensity recreation or spiritual uses and/or lowintensity extraction of plant and animal products and/or very sparse presence of infrastructure; key ecosystem functions such as carbon storage, biodiversity, and watershed protection and resilience expected to be very close to natural levels (excluding any effects from climate change) although some declines possible in the most sensitive elements (e.g., some high value hunted species).

Medium Forest Integrity (scores > 6.0 but <9.6) Interiors and natural edges of naturally regenerated forest ecosystems in blocks smaller than their natural extent but large enough to have some core areas free from strong anthropogenic edge effects (e.g., set-asides within forestry areas, fragmented protected areas), dominated by native species but substantially modified by humans through a diversity of processes that could include fragmentation, creation of edges and proximity to infrastructure, moderate or high levels of extraction of plant and animal products, significant timber removals, scattered stand-replacement events such as swidden and/or moderate changes to fire and hydrological regimes; key ecosystem functions such as carbon storage, biodiversity, watershed protection and resilience expected to be somewhat below natural levels (excluding any effects from climate change).

Low Forest Integrity (score ≤ 6.0): Diverse range of heavily modified and often internally fragmented ecosystems dominated by trees, including (i) naturally regenerated forests, either in the interior of blocks or at edges, that have experienced multiple strong human pressures, which may include frequent stand-replacing events, sufficient to greatly simplify the structure and species composition and possibly result in significant presence of non-native species, (ii) tree plantations and, (iii) agroforests; in all cases key ecosystem functions such as carbon storage, biodiversity, watershed protection and resilience expected to be well below natural levels (excluding any effects from climate change).

The numerical category boundaries were derived by inspecting FLII scores for a wide selection of benchmark locations whose forest integrity according to the category definitions was known to the authors, see text S6 and Table S4.

3.7. Protected areas analysis

Data on protected area location, boundary, and year of the inscription were obtained from the February 2018 World Database on Protected Areas [71]. Following similar global studies e.g. [72], we extracted protected areas from the WDPA database by selecting those areas that have a status of "designated", "inscribed", or "established", and were not designated as UNESCO Man and Biosphere Reserves. We included only protected areas with detailed geographic information in the database, excluding those represented as a point only. To assess the integrity of the protected forest, we extracted all 300m forest pixels that were at least 50% covered by a formally protected area and measured the average FLII score.

3.8. Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

4. Data availability

The authors declare that all data supporting the findings of this study are available at www.forestlandscapeintegrity.com. The datasets used to develop the Forest Landscape Integrity Index can be found at the following websites: tree cover and loss http://earthenginepartners.appspot.com/science-2013-global-forest, tree cover loss driver https://data.globalforestwatch.org/datasets/f2b7de1bdde04f7a9034ecb363d71f0e, potential forest cover https://data.globalforestwatch.org/datasets/potential-forestcoverage ESA-CCI Land Cover https://maps.elie.ucl.ac.be/CCI/viewer/index.php Open Street Maps https://www.openstreetmap.org, croplands https://lpdaac.usgs.gov/news/release-of-gfsad-30-meter-cropland-extent-products/, surface water https://globalsurface-water.appspot.com/, protected areas https://www.protectedplanet.net/en.

5. Acknowledgments

We thank Peter Potapov, Dmitry Aksenov, and Matthew Hansen for comments and advice. The research for this paper was in part funded by the John D. and Catherine T. MacArthur Foundation, Trillion Trees (a joint venture between BirdLife International, Wildlife Conservation Society, and WWF-UK), and other generous donors. It was also financially supported by UKAID from the UK government via the Forest Governance, Markets, and Climate Programme.

6. References

1. Seymour, F. & Harris, N. L. Reducing tropical deforestation. *Science* 365, 756-757 (2019).
2. Pearson, T. R. H., Brown, S., Murray, L. & Sidman, G. Greenhouse gas emissions from tropical forest degradation: an underestimated source. *Carbon Balance Manag*. 12, 3 (2017).
3. Moen, J. *et al.* Eye on the Taiga: removing global policy impediments to safeguard the boreal forest. *Conserv. Lett.* 7, 408-418 (2014).
4. Erb, K.-H. *et al.* Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature* 553, 73-76 (2018).
5. Scholes, R. *et al.* IPBES (2018): Summary for policymakers of the assessment report on land degradation and restoration of the Intergovernmental Science- Policy Platform on Biodiversity and Ecosystem Services. (2018).
6. Bridgewater, P., Kim, R. E. & Bosselmann, K. Ecological integrity: a relevant concept for international environmental law in the Anthropocene? *Yearb. Int. Environ. Law* 25, 61-78 (2014).
7. Parrish, J. D., Braun, D. P. & Unnasch, R. S. Are we conserving what we say we are? Measuring ecological integrity within protected areas. *Bioscience* 53, 851-860 (2003).
8. Cochrane, M. A. *et al.* Positive feedbacks in the fire dynamic of closed canopy tropical forests. *Science* 284, 1832-1835 (1999).
9. Erdozain, M. *et al.* Demand for provisioning ecosystem services as a driver of change in the Canadian boreal zone1. *Environ. Rev.* 27, 166-184 (2018).
10. Watson, J. E. *et al.* The exceptional value of intact forest ecosystems. *Nat. Ecol. Evol.* 1, 599-610 (2018).
11. Lewis, S. L. *et al.* Increasing carbon storage in intact African tropical forests. *Nature* 457, 1003 (2009).
12. Mello, K. D., Valente, R. A., Randhir, T. O. & Vettorazzi, C. A. Impacts of tropical forest cover on water quality in agricultural watersheds in southeastern Brazil. *Ecol. Indic.* 93, 1293-1301 (2018).
13. Garnett, S. T. *et al.* A spatial overview of the global importance of Indigenous lands for conservation. *Nat. Sustainability* 1, 369-374 (2018).
14. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* 320, 1444-1449 (2008).
15. Barlow, J. *et al.* Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* 535, 144 (2016).
16. Betts, M. G. *et al.* Global forest loss disproportionately erodes biodiversity in intact landscapes. *Nature* 547, 441 (2017).
17. Gibson, L. Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* 478, https://doi.org/10.1038/nature10425 (2011).
18. Di Marco, M., Ferrier, S., Harwood, T. D., Hoskins, A. J. & Watson, J. E. M. Wilderness areas halve the extinction risk of terrestrial biodiversity. *Nature* 573, 582-585 (2019).
19. Laurance, W. F. & Peres, C. A. Emerging threats to tropical forests. (University of Chicago Press, 2006).
20. Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A. & Schepaschenko, D. Boreal forest health and global change. *Science* 349, 819-822 (2015).
21. FAO. Assessing forest degradation. Towards the development of globally applicable guidelines. (FAO, Rome, 2011).
22. FAO. Global Forest Land-Use Change 1990-2005. (Food and Agriculture Organisation of the United Nations, Rome, 2012).
23. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* 342, 850-853 (2013).
24. Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* 361, 1108-1111 (2018).
25. Mitchell, A. L., Rosenqvist, A. & Mora, B. Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance Manag*. 12, 9 (2017).
26. Wang, Y. *et al.* Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote Sens. Environ.* 221, 474-488 (2019).
27. Venter, O. *et al.* Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* 7, 12558 (2016).
28. Shapiro, A. C., Aguilar-Amuchastegui, N., Hostert, P. & Bastin, J.-F. Using fragmentation to assess degradation of forest edges in Democratic Republic of Congo. *Carbon Balance Manag*. 11, 11 (2016).
29. Hansen, A. *et al.* Global humid tropics forest structural condition and forest structural integrity maps. *Sci. Data* 6, 232 (2019).
30. Williams, B. A. *et al.* Change in terrestrial human footprint drives continued loss of intact ecosystems. *One Earth* 3, 371-382 (2020).
31. Potapov, P. *et al.* The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv*. 3, https://doi.org/10.1126/ sciadv.1600821 (2017).
32. Watson, J. E. M. *et al.* Catastrophic declines in wilderness areas undermine global environment targets. *Curr. Biol.* 26, 2929-2934 (2016).
33. Olson, D. M. & Dinnerstein, E. G. The Global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. *Conserv. Biol.* 12, 502-515 (1998).

34. Pressey, R. L., Watts, M. E. & Barret, T. W. Is maximising protection the same as minimizing loss? Efficiency and retention as alternative measures of the effectiveness of proposed reserves. *Ecol. Lett.* 7, 1035-1046 (2004).

35. Maron, M., Simmonds, J. S. & Watson, J. E. M. Bold nature retention targets are essential for the global environment agenda. *Nat. Ecol. Evol.* 2, 1194-1195 (2018).

36. Grantham, H. S. *et al.* Spatial priorities for conserving the most intact biodiverse forests within Central Africa. *Environ. Res. Lett.* 15, 0940b0945 (2020).

37. Dudley, N. *et al.* The essential role of other effective area-based conservation measures in achieving big bold conservation targets. *Glob. Ecol. Conserv.* 15, e00424 (2018).

38. Laestadius, L. *et al.* Opportunities for forest landscape restoration. *Unasylva* 62, 238 (2011).

39. Edwards, D. P., Tobias, J. A., Sheil, D., Meijaard, E. & Laurance, W. F. Maintaining ecosystem function and services in logged tropical forests. *Trends Ecol. Evol.* 29, https://doi.org/10.1016/j.tree.2014.07.003 (2014).

40. Runting, R. K. *et al.* Larger gains from improved management over sparing-sharing for tropical forests. *Nat. Sustainability* 2, 53-61 (2019).

41. Chazdon, R. L. *et al.* A policy-driven knowledge agenda for global forest and landscape restoration. *Conserv. Lett.* 10, 125-132 (2017).

42. MacDicken, K. G. *et al.* Global progress toward sustainable forest management. *Ecol. Manag.* 352, 47-56 (2015).

43. Ghazoul, J. & Chazdon, R. Degradation and recovery in changing forest landscapes: a multiscale conceptual framework. *Annu. Rev. Environ. Resour.* 42, 161-188 (2017).

44. Strassburg, B. B. N. *et al.* Strategic approaches to restoring ecosystems can triple conservation gains and halve costs. *Nat. Ecol. Evolution* 3, 62-70 (2019).

45. Pandit, R. *et al.* Summary for policymakers of the assessment report on land degradation and restoration of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (2018).

46. Osuri, A. M. *et al.* Contrasting effects of defaunation on aboveground carbon storage across the global tropics. *Nat. Commun.* 7, 11351 (2016).

47. Mackey, B. Policy options for the world's primary forests in multilateral environmental agreements. *Conserv. Lett.* 8, https://doi.org/10.1111/ conl.12120 (2015).

48. Butchart, S. H. M., Di Marco, M. & Watson, J. E. M. Formulating smart commitments on biodiversity: lessons from the Aichi targets. *Conserv. Lett.* 9, 457-468 (2016).

49. Watson, J. E. *et al.* Set a global target for ecosystems. (2020).

50. Stuart, S. N. *et al.* Clarifying the key biodiversity areas partnership and programme. *Biodivers. Conserv.* 27, 791-793 (2018).

51. Jones, K. R. *et al.* One-third of global protected land is under intense human pressure. *Science* 360, 788-791 (2018).

52. Fa, J. E. *et al.* Importance of Indigenous Peoples' lands for the conservation of Intact Forest Landscapes. *Front. Ecol. Environ.* 18, 135-140 (2020).

53. Barrington-Leigh, C. & Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS ONE* 12, e0180698 (2017).

54. Hughes, A. C. Have Indo-Malaysian forests reached the end of the road? *Biol. Conserv.* 223, 129-137 (2018).

55. Pasher, J., Seed, E. & Duffe, J. Development of boreal ecosystem anthropogenic disturbance layers for Canada based on 2008 to 2010 Landsat imagery. *Can. J. Remote Sens.* 39, 42-58 (2013).

56. Mackenzie, C. A. & Hartter, J. Demand and proximity: drivers of illegal forest resource extraction. *Oryx* 47, 288-297 (2013).

57. Lim, C. L., Prescott, G. W., De Alban, J. D. T., Ziegler, A. D. & Webb, E. L. Untangling the proximate causes and underlying drivers of deforestation and forest degradation in Myanmar. *Conserv. Biol.* 31, 1362-1372 (2017).

58. Erb, K.-H. *et al.* Land management: data availability and process understanding for global change studies. *Glob. Change Biol.* 23, 512-533 (2016).

59. Harris, N., E., Goldman, L. & Gibbes, S. Spatial Database of Planted Trees (SDPT) Version 1.0. (World Resources Institute, Washington, DC, 2018).

60. Gorelick, N. *et al.* Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18-27 (2017).

61. Heino, M. *et al.* Forest loss in protected areas and intact forest landscapes: a global analysis. *PLoS ONE* 10, e0138918 (2015).

62. Venter, O. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* 7, https://doi. org/10.1038/ncomms12558 (2016).

63. Laurance, W. F. *et al.* Ecosystem decay of Amazonian forest fragments: a 22 year investigation. *Conserv. Biol.* 16, 605-618 (2002).

64. Cochrane, M. A. & Laurance, W. F. Fire as a large-scale edge effect in Amazonian forests. *J. Tropical Ecol.* 18, 311-325 (2002).

65. Peres, C. A., Emilio, T., Schietti, J., Desmoulière, S. J. M. & Levi, T. Dispersal limitation induces long-term biomass collapse in overhunted Amazonian forests. *Proc. Natl Acad. Sci. USA* 113, 892-897 (2016).

66. Zimmerman, B. L. & Kormos, C. F. Prospects for sustainable logging in tropical forests. *Bioscience* 62, 479-487 (2012).

67. Chaplin-Kramer, R. *et al.* Degradation in carbon stocks near tropical forest edges. *Nat. Commun.* 6, 10158 (2015).

68. Maisels, F. *et al.* Devastating decline of forest elephants in Central Africa. *PLoS ONE* 8, e59469 (2013).

69. Berzaghi, F. *et al.* Carbon stocks in central African forests enhanced by elephant disturbance. *Nat. Geosci.* 12, 725-729 (2019).

70. Beyer, H. L., Venter, O., Grantham, H. S. & Watson, J. E. M. Substantial losses in ecoregion intactness highlight urgency of globally coordinated action. *Conserv. Lett.*, e12592, https://doi.org/10.1111/conl.12692.

71. IUCN, U.-W. a. Protected Planet: The World Database on Protected Areas (WDPA). (UNEP-WCMC, Cambridge, UK, 2018).

72. Butchart, S. H. M. *et al.* Shortfalls and solutions for meeting national and global conservation area targets. *Conserv. Lett.* 8, 329-337 (2015).

Punta Begoñako grafitietan garbiketa kimiko selektiboa egin behar da

Punta Begoñako Galerien eraikin historikoa zaharberritzeko lanak abiatu nahi dira. Lan horretan, beste hainbat urratsen artean, denborarekin egin diren grafitiak kendu nahi dituzte. EHUko kimikari-talde batek grafitien pigmentuak eta aglutinatzaileak aztertzeko proiektu bat egin du, lanari ekin baino lehen garbiketa-metodo egokienak kimikoki bilatzeko. Zaharberritze historikoak baliabideak behar ditu, prozesuak ez baitira sinpleak izaten.

Punta Begoñako Galeriak Getxon daude, Bizkaian, itsasoaren ondoan. Denbora luzea eman dute utzita, eta grafiti asko egin dira paretetan. Galeriak zaharberritzearen ikuspuntutik, horrek esan nahi du pigmentuen geruza asko daudela paretetan bata bestearen gainean. Gela nagusian, gainera, grafitiak jatorrizko margolan batzuen gainean daude. Hain zuzen ere, zaharberritze-prozesuaren zailtasunetako bat da margolanak hondatu edo galdu gabe kentzea grafitiak.

Askotan garbiketa-prozesu horiek saiakuntzan eta errakuntzan oinarritzen dira, baina kasu honetan beste estrategia bat proposatu dute adituek. Ikerketa batean, grafitien pigmentuak eta aglutinatzaileak aztertu dira. Batetik, haien osagaiak identifikatu dituzte. Bestetik, denboraren efektua ere aztertu dute ikertzaileek, inguruak pigmentuetan kutsatzaile atmosferiko batzuk txertatu eta zenbait korrosio-prozesu eragin baititu. Hori guztia argituta, ikertzaileen lanak oinarri bat ezarri du metodologia kimiko bat prestatzeko, garbiketari ekiteko.

Ikerketaren deskribapena

Emaitzetatik argi geratu da Punta Begoñako Galerietako grafitiak 15 esprai ezberdinekin egin direla, nahiz eta esprai komertzialen osagaien informazioa ez den publikoa. Informazio hau eskuratzeko, kimikari analitikoek ohiko bi teknika erabili dituzte: Raman espektroskopia eta X izpien fluoreszentzia.

Lehenengo teknikak pigmentuen molekulak identifikatzen ditu, argi infragorriaren bitartez. Osagai nagusiak topatuta, ikertzaileek jakin dute esprai komertzial ohikoenak erabili izan direla grafitietan. Haietaz gain, hirugarren markako espraien pigmentuak identifikatu dituzte. Bigarren teknikak metal-atomoak identifikatzen ditu, eta, analisi horretan, espero zitekeen emaitza bat nabarmena izan da; Punta Begoñako grafitien pigmentuetan titanioa da nagusi, titanio oxido moduan. Esprai erabilienen pigmentuetan da ohikoa titanio oxidoa.

«Punta Begoñako grafitiak garbitzea zaila da, jatorrizko margolanak hondatu edo galdu gabe»

Kolore bakoitza metal bati lotuta dago. Hala ere, oso merkatu aldakorra da; etengabe ari dira espraietan osagai berriak merkaturatzen, eta horrek zaildu egiten du pigmentuen identifikazioa. Gainera, itsasoak berak askatutako aerosolek elementu gehiago utzi dituzte bai grafitietan bai eta Punta Begoñako paretetan ere.

Punta Begoñako Galerietako gela nagusiko grafitiek pigmentu-geruza asko dituzte, osagai organiko eta inorganiko asko, eta, beraz, garbiketa-lana ezin da erraza izan. Alde batetik, ikertzaileek proposatu dute urrats askoko garbiketa-prozesu bat egitea, geruza bakoitza kentzeko berezko produktu bat erabilita. Bestalde, eskuratu duten informazioa eskuragarri jarri dute. Eta esprai komertzialen osagaien datu-base ireki baten beharra azpimarratu dute.

InpAkta 00 (2023) 59-73
https://doi.org/10.3390/app11188640

Graffiti characterization prior to intervention in the Punta Begoña galleries (Getxo, north of Spain): Raman and XRF spectroscopy in the service of restoration

Idoia Etxebarria[1], Nagore Prieto-Taboada[1,*], Estibaliz Lama[2], Gorka Arana[1], María Dolores Rodríguez-Laso[2] and Juan Manuel Madariaga[1,3]

[1] Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country UPV/EHU; [2] Department of Painting, Faculty of Fine Arts, University of the Basque Country UPV/EHU; [3]Unesco Chair of Cultural Landscapes and Heritage, University of the Basque Country UPV/EHU

ABSTRACT: The Historical and Cultural Heritage of Punta Begoña Galleries in Getxo (Bizkaia, North of Spain) are currently in restoration after being abandoned for years. For that reason, many graffiti, which directly affect the wall paintings, appear on most of their walls. Moreover, several graffiti overlap each other, which makes their removal more difficult. For all these reasons, the chemical characterization of these pigments is a priority to optimize the cleaning and consolidation treatments of wall paintings. That being the case, an analysis based on Raman spectroscopy and X-ray fluorescence was carried out to obtain information to help conservators remove the graffiti without damaging the mural paintings and the support. Nevertheless, the first step, using X-ray fluorescence and Raman spectroscopy, involved the need for a database to compare the results and identify the compounds. Thus, different commercial inks were elementally and molecularly characterized to complete the existing databases. After this, an analysis of the inks was carried out that noted the presence of several organic pigments, such as phthalocyanines. Inorganic pigments such as titanium oxide were identified as well. After the analysis, the selection of the best removal process could be carried out to provide the most effective treatment, avoiding the "trial-and-error" classical practice.

1. Introduction

The Historical and Cultural Heritage of Punta Begoña Galleries in Getxo (Bizkaia, North of Spain) were built in 1918 as a sign of the economic power of the businessman Horacio Etxebarrieta, with the aim of restraining the hillside and creating a private leisure area. The building featured several

* **Corresponding author / Harremanetan jartzeko:** Nagore Prieto-Taboada. Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country UPV/EHU, P.O. Box 644, 48080 Bilbao, Basque Country, Spain. e-mail: nagore.prieto@ehu.eus. **Contributed by** Olatz Zuloaga **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Etxebarria, I. *et al.* «Graffiti Characterization Prior to Intervention in the Punta Begoña Galleries (Getxo, North of Spain): Raman and XRF Spectroscopy in the Service of Restoration», *Applied Sciences* | (2021)11:8640 (https://doi.org/10.3390/app11188640)

technological innovations, such as the use of reinforced concrete and other cutting-edge materials for that time [1].

Over the years, the Galleries were abandoned, but recently have been recovered through a multidisciplinary project that aims to restore value to building and the history that surrounds it [2]. The hall room has special relevance as a place of celebrations and events, and currently is in a precarious state of conservation. This hall presents paintings with Francoist iconography related to the wars of the 20th century, but recent studies indicate the presence of original underlying paintings [1]. However, the state of neglect in recent years has led to the presence of numerous graffiti that directly affect the wall paintings and, more importantly, put at risk the conservation of the original mural paintings [1].

For that reason, it is very important to remove the graffiti in a scientific way, avoiding the usual "trial-and-error" method in conservation works, to minimize the impact on the support on which they are found. In fact, the stopping of the attacks and damage caused by graffiti on cultural heritage is currently a priority in conservation efforts in general [3-5]. In addition, graffiti generates a vision of abandonment in these places of cultural interest, and usually aid their deterioration due to the interaction with atmospheric agents [6].

Thus, a chemical characterization of the graffiti seems to be essential to propose the best cleaning methodology. The characterization of the pigments is of great help in directing the cleaning and elimination processes for these layers, as well as in evaluating their impact on the compositional materials [7-9]. These paints are generally composed of binders, pigments, fillers, and additives [10].

The chemical characterization of pigments used in graffiti is also crucial in general for the conservation and restoration of urban art, which has increased in recent years in our cities, and not only for the recovery of degraded heritage by anthropogenic factors (vandalism). Two of the most interesting techniques are Raman spectroscopy, for the identification of the molecules composing the different pigments, and X-ray fluorescence (XRF), for the identification of elements present in them [11-14]. Both techniques allow the understanding of the ink's composition in a complementary way.

The first step for the characterization of pigments by Raman spectroscopy involves the need for a valuable database to compare the results. This fact becomes more important when bearing in mind the quick development of the graffiti industry and the associated secrecy. Thus, complete and updated databases are required. In the literature, it is possible to find free databases [8, 9, 11, 15, 16], but most of them are not open-access, which limits the possibility of exchanging the information and joining forces. Regardless, there are some interesting open-access databases, such as the Royal Institute for Cultural Heritage database (KIK-IRPA) [15] or the spectral database of the Infrared and Raman Users Group (IRUG) [16], although these reference spectra have been limited to pigments and only a certain number of them, not to specific commercial sprays, which contain a mixture of compounds.

Thus, commercial spray databases would be very relevant, considering aerosol paint manufacturers are reluctant to offer the consumer the overall composition of paints [8,9], usually because each producer keeps its own industrial manufacturing secrets for materials and techniques. For this reason, they only provide the most general composition of their sprays [17]. Nevertheless, if the artists and their most-used sprays are known, the analysis and conservation of their artwork could be easier, or even could help to determine the real authorship of the artwork [9]. It is therefore necessary to create a database containing as many commercial aerosols as possible before beginning any analysis. Taking this into account, it is not only important to have a database with the Raman spectra of a specific pigment, but also relating it to the brand or model of spray, which could provide relevant information and solve the problem of lack of transparency on the part of the industry. Moreover, the elemental information would give important information about the composition of the different brands; for example, allowing their identification. Therefore,

the databases of graffiti would be completed and updated with all this information.

On the other hand, the rapid evolution of this industry generates the launch of new sprays and components very often, which could render the databases obsolete. That is why databases, especially for these types of compounds, should be published as open access in order to share, complete, and improve the information available. In this way, the industrial secret problem could be minimized or even settled.

All things considered, the aim of this work was characterizing the graffiti observed in the main hall of the Punta Begoña Galleries by means of Raman spectroscopy and XRF analysis to help restorers to design the most adequate procedure to carry out restoration works. First, 15 different commercial sprays were measured to enlarge the existing open-access databases found in the literature.

Figure 1: (Top) photograph of the main hall of the Punta Begoña Galleries showing the graffiti covering the entire surface of the marble and mural paintings, up to a height of approximately 2 m; (bottom left) mock-ups with the applied commercial sprays; (bottom right) example of the sampling of the inks in the main hall

2. Materials and Methods

2.1. Commercial Sprays Analyses: Database

Fifteen commercial sprays were selected while considering the colours observed in Punta Begoña's main hall. The selected commercial sprays were different brands. From the Titanlux brand (Tintalux, Barcelona), three sprays of the "spray colour" series were selected: Orange (554), White (566) and Yellow (529). From the Montana brand (Montana Colours, Barcelona) Hardcore series, five sprays were chosen: Vivid Red (RV-3001), Arctic Blue (RV-29), Dark Blue (RV-5005), Druid Brown (RV-246), and Cologno N. Green (RV-237). Finally, from the Decor Pintura brand (ZTHOME, Logroño) seven sprays were analysed: Medium Yellow (ZT107), Light Grey (ZT113), May Green (ZT129), Light Brown (ZT120), Violet (ZT131), Fuchsia (ZT134) and Bordeaux (ZT148).

The sprays were applied on ceramics supports specially made for assays in this project [18]. The mock-ups were 8 × 10 cm in size (Figure 1). They were left to dry for 20 days and then collected by scalpel and measured in the same conditions as the inks of the hall room.

2.2. Sampling

The graffiti cover the entire lower part of the main hall. Up to about 1.5 m, the wall is covered with decorative marble, and from that height and above the original mural paintings can be seen. At first, an attempt was made to carry out in situ analysis by portable Raman spectrometer (innoRam laser at 785 nm excitation, BWTek, Newark, USA) and XRF instrumentation (XMET5100, Oxford Instruments, Abingdon, Oxfordshire, UK). However, the results were poor due to dirt, patina, and fluorescence. For this reason, it was decided to sample the graffiti.

The inks were sampled in the marble area to avoid damaging the paintings, which are in a very precarious state of conservation. For this purpose, and considering that sampling is very restricted, the area with the greatest accumulation of graffiti was chosen to chronologically collect the different types of ink present on the walls and also see the effect of their overlapping. The building has been abandoned for decades, so the results give information for all periods. This could be relevant, due to the changes in the composition of sprays over the years.

The layers were scraped by scalpel and collected on different papers to avoid cross-contamination, in an attempt to discriminate between colours or intermediate phases, and disregarding the areas that resembled these intermediate mixtures. As mentioned, the inks were collected consecutively at the same point, the first one being the most modern, and the last one the oldest. The colours collected were: (T1) light red, (T2) light green, (T3) dark yellow, (T4) light blue, (T5) light orange, (T6) dark blue, (T7) white, (T8) bone white, (T9) ochre, (T10) dark green, (T11) purple, and (T12) light yellow. Special care was taken in the selection of the specific colours of each ink, without mixing several of them to facilitate their characterisation (Figure 1). Samples were stored in 1.5 mL Eppendorf tubes until analysis. They were then stored in labelled airtight zip bags.

2.3. Instrumentation

The 12 inks collected from the Punta Begoña main hall and the 15 commercial sprays were analysed by Raman spectroscopy and X-ray fluorescence (XRF).

A Renishaw InVia high-resolution micro-Raman spectrometer (Renishaw, UK) was used. The instrument was coupled to a microscope (Leica, Germany) and the objectives used for the measurements were 20×, 50×, and 100× to provide a better focus on the surface. Raman spectra were acquired, in general, between 100-3000 cm^{-1}, with a spectral resolution of 1 cm^{-1}. The lasers used for the analyses were 532 nm and 785 nm, with a power of 45 mW, although a maximum of 10% of this power was used in the analyses. A CCD detector cooled by the Peltier effect was used, and the program used by the equipment was the Wire 4.2 software package (Renishaw, Gloucestershire, UK). The treatment of the Raman results was performed by the Omnic 7.2 software (Nicolet).

For elementary analyses, an X-ray fluorescence M4 Tornado (Bruker Nano GmbH, Germany) spectrometer was used. The instrument implemented two Rh tubes mounted on a mechanical collimator (lateral resolution 1 mm) and a policapillary optic (lateral resolution 25 μm). The X-ray tube worked at 50 kV and 700 μA during the measurements, and 100 s (lifetime) were considered for each spectral acquisition. The measurements were made under vacuum (20 mbar). To achieve the vacuum, an MV10N VARIO-B diaphragm pump (Vaccubrand, Germany) was used. Relative composition information is provided, using a semiquantitative analysis carried out by the software based on fundamental parameters. The spectral data acquisition and treatment was performed using the M4 software from Bruker.

3. Results

3.1. Commercial Sprays

The identification of Raman spectra is complex due to the large number of bands, the small differences between different compounds of a similar nature, and problems with obtaining quality spectra. Despite this, it was possible to identify the main pigments of all the commercial sprays by means of Raman spectroscopy, and even some of the bands related to the binders. Moreover, the elemental analysis, which is not so often used in spray characterizations, provided relevant information to better understand their composition.

The results of the XRF analysis of the commercial sprays are collected in Table 1, and molecular information and the spectra obtained by Raman spectroscopy are collected in Table 2 and Figures 2-4. As can be seen in Table 1, most of the commercial sprays had titanium as a major element. This was usually present as titanium oxide in rutile form, a common compound used in sprays as an extenders and opacifier of paints [7, 8].

However, it was evident that this pattern is not present in the "Decor" brand, in which titanium was not a major element. In fact, in general, in this brand the elemental composition of the spray was more complex. This could be related to the quality of the sprays, since the "Titan" and "Montana" are sprays generally used in street art [6]. Considering the above, the number of major elements was lower, which could indicate: (1) a simpler composition of the mixture; or (2) the use

of organic compounds that cannot be analysed by XRF were used in conjunction with rutile as an additive. However, regarding the molecular analysis, it was striking that the identification of rutile [16], despite being the major component in many sprays, was only identified when it was used as pigment as in the case of "White 566" and "Light Grey ZT113", and in a pigment mixture in "Arctic Blue", in which the rutile was mixed with the PB15 phthalocyanine blue [15]. Another remarkable element observed in the XRF analysis was iron, present mostly in sprays of the red and yellow palette, and especially in the brown colours "Druid Brown RV-246" and "Light Brown ZT120". In these sprays, Raman spectroscopy revealed a mixture ofiron oxidem, hematite, and iron hydroxide, goethite [16]. Finally, the presence of copper was also significant due to its presence as major element in the "Dark Blue RV-5005" spray, which could indicate the use of a pigment based on this element that is usual in blue colours [19]. As was presumable by the elemental analysis results, the pigment PB15, a copper pthalocyanine, was identified by Raman spectroscopy [15].

Table 1
Elements identified in the commercial sprays. The major elements were considered up to 10% of relative presence. The elements with less than 1% were not included. The elements were ordered by taking into account their relative presence

Brand	Pigment	Major elements	Minor elements
Titan	Orange (554)	Ti, Fe, Cl	Zn, Pb, Co, Zr, Ca
	White (566)	Ti	—
	Yellow (529)	Ti, Fe	Ca, Zr, Co, Cl, Zn
Montana	Vivid Red (RV-3001)	Ti, Fe	Ca, Co, Zr, Zn
	Artic Blue (RV-29)	Ti	—
	Dark Blue (RV-5005)	Ti, Cu	Si, Ca, Fe
	Druid Brown (RV-246)	Ti, Fe	Zr
	Cologno N. Green (RV-237)	Ti, Fe	Cl, Ca, Cu, Si, Zr, Co
Decor Pintura	Medium Yellow (ZT107)	Bi, V	Zr, Co, Zn, Ca, Ti, Cl
	Light Grey (ZT113)	Ti	Fe, Si, Zr, Ca
	May Green (ZT129)	Cl, Cu, Zr, Ca	Ti, Fe, Co
	Light Brown (ZT120)	Fe	Zr, Si, Zn, Ca, Co
	Violet (ZT131)	Ti, Fe, Zr	Ca, Co, Zn
	Fuchsia (ZT134)	Ti	Fe, Zr, Cl, Co, Ca, Cr
	Bordeaux (ZT148)	Cl, Fe, Ca, Si, Zr	Co, Ti, Zn

Table 2
Compounds identified by means of Raman spectroscopy in the commercial sprays

Brand	Pigment	Composition	Raman bands
Titan	Orange (554)	Pyrazoloquinazolone (PO67)	121, 226, 407, 546, 627, 652, 666, 738, 1012s, 1117, 1154, 1166, 1216, 1256s, 1274, 1298, 1322, 1340, 1374, 1413, 1532, 1558s, 1577, 1603s cm^{-1}
	White (566)	Titanium oxide (Rutile)	143, 234, 449s, 612s cm^{-1}
		Alkyd binder Binder	1044, 1168, 1302, 1447, 1582, 1602, 1725 cm^{-1}
	Yellow (529)	Diarylide (PY13)	620, 656, 1146, 1259s, 1288s, 1401s, 1450, 1493, 1526, 1600vs cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1004,1044, 1166, 1727 cm^{-1}
Montana	Vivid Red (RV-3001)	Naphthol AS (PR170)	112, 193, 208, 267, 292, 335, 364, 423, 478, 494, 511, 539, 574, 611, 732s, 778, 806, 866, 963s, 1013, 1043, 1117, 1141, 1167s, 1208, 1245, 1288s, 1334, 1364vs,1386, 1423, 1454, 1489, 1513s, 1552, 1606vs cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1002, 1043, 1725 cm^{-1}
	Artic Blue (RV-29)	Phthalocyanine (PB15)	175, 234, 258, 484, 681s,748s, 782, 833, 849, 955, 1110, 1145, 1196, 1218, 1309, 1344s, 1453, 1530vs, 1602 cm^{-1}
		Titanium oxide (Rutile)	142, 234, 447s, 610s cm^{-1}
		Alkyd binder	1008, 1042, 1731 cm^{-1}
	Dark Blue (RV-5005)	Phthalocyanine (PB15)	176, 235, 259, 483, 483, 595, 681s, 719, 747, 832, 847, 954, 1038, 1108, 1144, 1194, 1216, 1307, 1342s, 1412, 1452s, 1470, 1483, 1529vs, 1590 cm^{-1}
	Druid Brown (RV-246)	Hematite	226, 244, 292s, 404s, 484, 610 cm^{-1}
		Goethite	392, 549 cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	652, 1004, 1040, 1578, 1601 cm^{-1}
	Cologno N. Green (RV-237)	Phthalocyanine (PG7)	194, 234, 266, 598, 641, 684, 740, 775, 818, 951, 975, 1212, 1261, 1282, 1292, 1337, 1504, 1537, 1590 cm^{-1}

Brand	Pigment	Composition	Raman bands
Decor Pintura	Medium Yellow (ZT107)	Acetoacetic arylide (PY74)	186, 329, 363, 406, 648, 803, 829, 923, 1068, 1090, 1161, 1171, 1265, 1299, 1327, 1354, 1406, 1440, 1491, 1514, 1563, 1595, 1669 cm^{-1}
		Alkyd binder	726, 1043, 1727 cm^{-1}
	Light Grey (ZT113)	Titanium white	144, 233, 444s, 611s cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1005, 1043, 1303, 1439, 1602, 1728 cm^{-1}
	May Green (ZT129)	Phthalocyanine (PG7)	148, 165, 198, 223, 266, 291, 333, 347, 546, 594, 643, 685s, 740s, 776s, 817, 956, 979, 1083, 1213s, 1282s, 1338s, 1388, 1445, 1538vs cm^{-1}
	Light Brown (ZT120)	Hematite	226, 246, 293, 411, 610, 1300 cm^{-1}
		Goethite	387, 549 cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1005, 1042, 1444, 1583, 1602, 1724 cm^{-1}
	Violet (ZT131)	Dioxazine (PV23)	317, 416, 484, 528 s, 591, 672, 724, 748, 921, 934, 990, 1108, 1122, 1131, 1152, 1167, 1208 s, 1258 s, 1334 s, 1348s, 1394 vs, 1433 s, 1445, 1593 cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1003, 1042 cm^{-1}
	Fuchsia (ZT134)	Dioxazine (PV23)	317, 485, 529, 592, 614, 681, 726, 746, 812, 922, 935, 1110, 1133, 1165, 1208, 1260, 1335, 1349, 1394s, 1433, 1447, 1461, 1581s cm^{-1}
		Alkyd binder, most likely CHS-Alkyd S 471	1004, 1042, 1601, 1726 cm^{-1}
		Unidentified	347, 442, 1062, 1231, 1284, 1359, 1484, 1553 cm^{-1}
	Bordeaux (ZT148)	Naphthol AS (PR112)	348, 431, 529, 575, 620, 682, 720, 747, 814, 967, 1063, 1163, 1231, 1245, 1358, 1375, 1393, 1403, 1485, 1555, 1581 cm^{-1}
		Alkyd binder	651, 1004, 1729 cm^{-1}

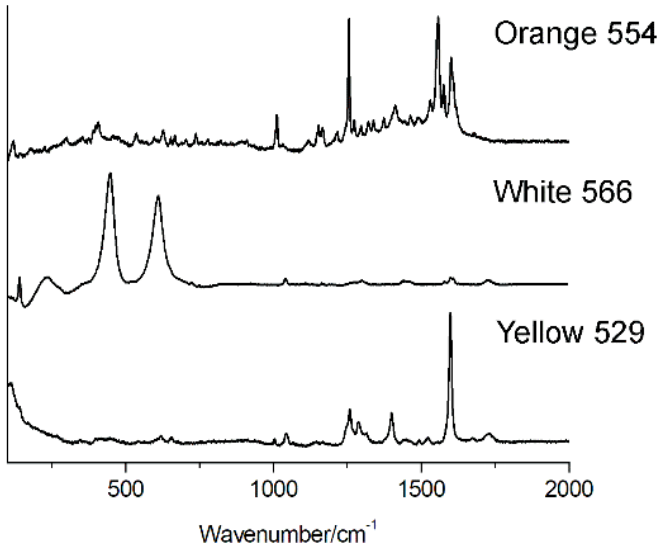


Figure 2: Raman spectra of the commercial sprays of the Titan brand

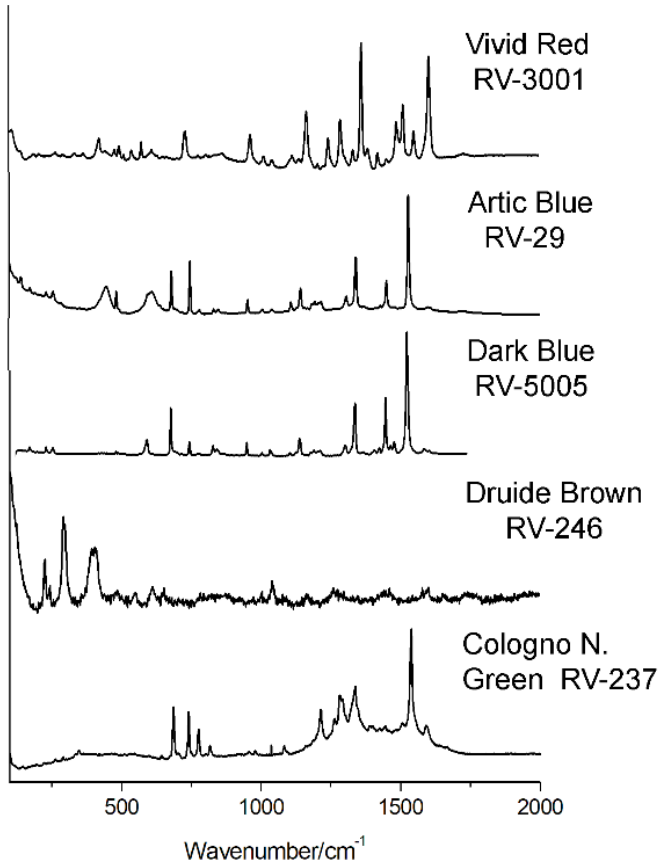


Figure 3: Raman spectra of the commercial sprays of the Montana brand

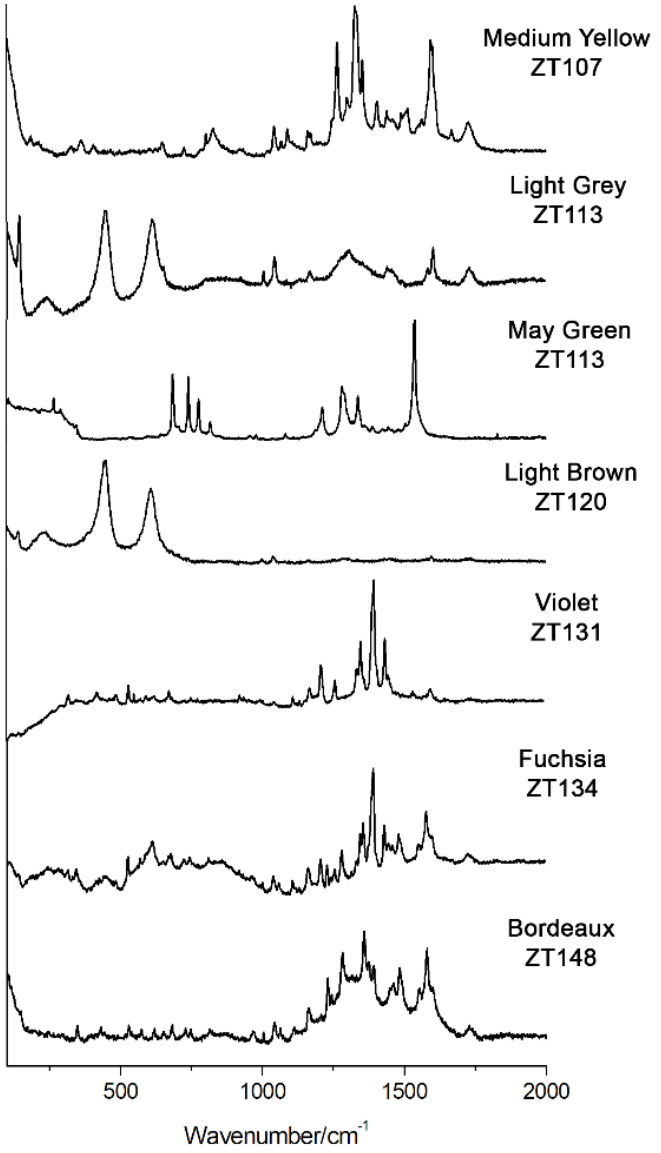


Figure 4: Raman spectra of the commercial sprays of the Decor brand

Regarding the "Decor" sprays, the composition changed considerably. As was mentioned previously, titanium was not present as a major element in all of these sprays, and other ones appeared, most likely indicating the use of a different type of formulation than in the previous brands, and even between different colours. In this sense, this brand had also more minor elements, indicating a more complex mixture. Regardless, the elemental analysis helped us to better understand the composition of the sprays. For example, in "Medium Yellow ZT107", the major elements were bismuth and vanadium, which could indicate

the use of bismuth vanadate ($BiVO_4$), a common pigment used in the last decades [8]. However, in the Raman spectroscopy analysis, only acetoacetic arylide PY74, with the formula $C_{18}H_{18}N_4O_6$, was identified as a pigment [15]. Considering the relative presence of the bismuth and vanadium, the mixture of both yellow pigments (PY74 and $BiVO_4$) in this spray seems plausible. In the same way, in the cases of "May Green ZT129" and "Cologno N. Green RV-237", the presence of copper and chlorine as major elements could indicate the use of phthalocyanine green, which includes both elements in their structure. In addition, the PG7 pigment effectively was identified in both cases [15]. Regarding this point, it is necessary to remark that in the case of "Cologno N. Green RV-237", three Raman bands were observed at 1261, 1504, and 1590 cm^{-1}; these bands are not found in the literature [15,16,20]. In fact, some discrepancies were observed in the literature, but the spectrum collected in the IRUG database [16] finally presented broad bands that could contain the bands that appear in this study, thus, the resolution of the obtained spectra most likely allowed us to obtain more than one band instead a broad one. Continuing with the chlorinated pigments, in "Bordeaux ZT148", chlorine was identified as major element due to the presence of naphthol AS PR112 pigment, a chlorinated pigment with the formula $C_{24}H_{16}Cl_3N_3O_2$ [15], but the presence of iron was also remarkable. Considering that the spray was red, the use of iron oxides mixed with the mentioned organic pigment was probable. The same case occurred with the "Vivid Red RV-3001" pigment, in which iron also was identified as major element, but when using Raman spectroscopy, another naphthol AS was identified as a pigment, in this case PR170 ($C_{26}H_{22}N_4O_4$), without chlorine [15]. Another chlorinated pigment was assumed in the "Orange 554" spray by XRF, and was identified as pyrazoloquinazolone PO67 ($C_{17}H_{11}ClN_6O_3$) pigment by Raman spectroscopy [15]. The same occurred for "Yellow 529", which was composed of diarylide PY13 pigment, with the formula $C_{36}H_{34}Cl_2N_6O_4$ [15]. Finally, the "Violet ZT131" and "Fuchsia ZT134" were very similar, as they were composed of the dioxazine PV23 pigment ($C_{34}H_{22}Cl_2N_4O_2$) [15]. However, in the case of the latter, other bands were observed by Raman spectroscopy, but it was not possible to clearly identify the compound related to them. The presence of a monoazo or diazo red pigment could explain the bands 347, 1359, 1484, and 1553 cm^{-1}, but with an incomplete identification. In the XRF, it was also clear that the composition of both sprays was different even though the same pigment was present. In this sense, violet presented zinc, and fuchsia presented chlorine and chrome.

The binder used could be identified in many of the sprays. In all cases, they appeared to be alkyd binders [6, 21, 22], among the most common binders used in art due to their low cost, fast drying, and good optical properties. In this sense, the most repeated binder identified was Alkyd S471, with a 1042 cm^{-1} Raman band related to phthalate and a 1004 cm^{-1} band related to isophthalate [22]. Nevertheless, other alkyd binders were observed, with very similar Raman spectrum to S471 but slightly different, with the most relevant modifications being the appearance or disappearance of one of these two mentioned characteristic bands (1004 or 1042 cm^{-1}), or the movement of the 1725-1732 cm^{-1} broad band, but related in any case to alkyd binders [22].

Taking all this into account, all the selected sprays could be identified, and although the complete composition was not achieved, information for pigments and binders was obtained, as well as for other additional compounds. This data increases the public database, with additional information to the usual information available as the brand or elemental composition, emphasizing the need to compile databases of real sprays and not just pure pigments. Moreover, all of this was important to do in a total open-access way, so the spectra of these studies are accessible in the Supplementary Materials (Table S1). Furthermore, as occurred with the PG7 pigment, it is important to provide not only the bands, but also the spectra, since, given the complexity of the spectra, discrepancies can be found in the literature, and without the raw spectra, the comparison is more difficult.

3.2. Inks Samples from the Main Hall

Regarding the sampled inks from the main hall, the analysis by XRF was more complicated due to the effect of the environment, as can be seen in Table 3. In

the first layers, it was possible to observe as major elements sulphur and calcium, the relative concentration of which decreased with the depth. In fact, sulphur in the inner layers appeared as a minor element. This was due to the action of atmospheric acid gases that have considerably affected the building [23,24]. In this sense, it is remarkable that the inks acted as a protective layer, minimizing the impact of the atmospheric SO_2 on the original material, the result of wet or dry deposition due to the marine aerosol and usual fogs in the area, in contrast to the supposed increase of the degradation that these inks generate in the original materials [6]. This protection most likely will be the consequence of the high number of ink layers. Leaving aside this effect, elements such as silicon and potassium could be also related to the particulate matter, due to the marine aerosol and sand from the near beach [1, 18, 24]. Taking all of these in consideration, titanium again seems to be a relevant element in the spray's compositions. However, once again, rutile was observed by Raman spectroscopy only when it was used as pigment and not when used as additive, possibly due to the quantity present [16].

Table 3
Elements identified in the sample's colours. The major elements were considered up to 10% of relative presence. The elements with less than 1% were not included. The elements were ordered by taking into account the relative presence

Name	Major elements	Minor elements
(T1) Light Red	Ca, S	Fe, Ti, Pb
(T2) Light Green	Ti, Ca	S, Fe, Pb, Si, Zr
(T3) Dark Yellow	Ca, Fe, S	Ti
(T4) Light Blue	Ti, Ca, Fe	S, Pb, Bi, Si, K
(T5) Light Orange	Ca, Fe, Ti	S, Pb, K, Si
(T6) Dark Blue	Ti, S, Ca	Pb, Fe, Cu
(T7) White	Ti, Ca	Fe, Si, S
(T8) Bone White	Ti, Fe	Cu, Ca, S, Zr
(T9) Ochre	Fe, Ca, S, Ti	Cu, Pb
(T10) Dark Green	Ca, Ti, Fe, S	Cl, Si Cu, Pb, K
(T11) Purple	Ti, Fe, Ca	S, Bi, Si, Mn, Zn
(T12) Light Yellow	Ti, Fe	Ca, S, Si, As

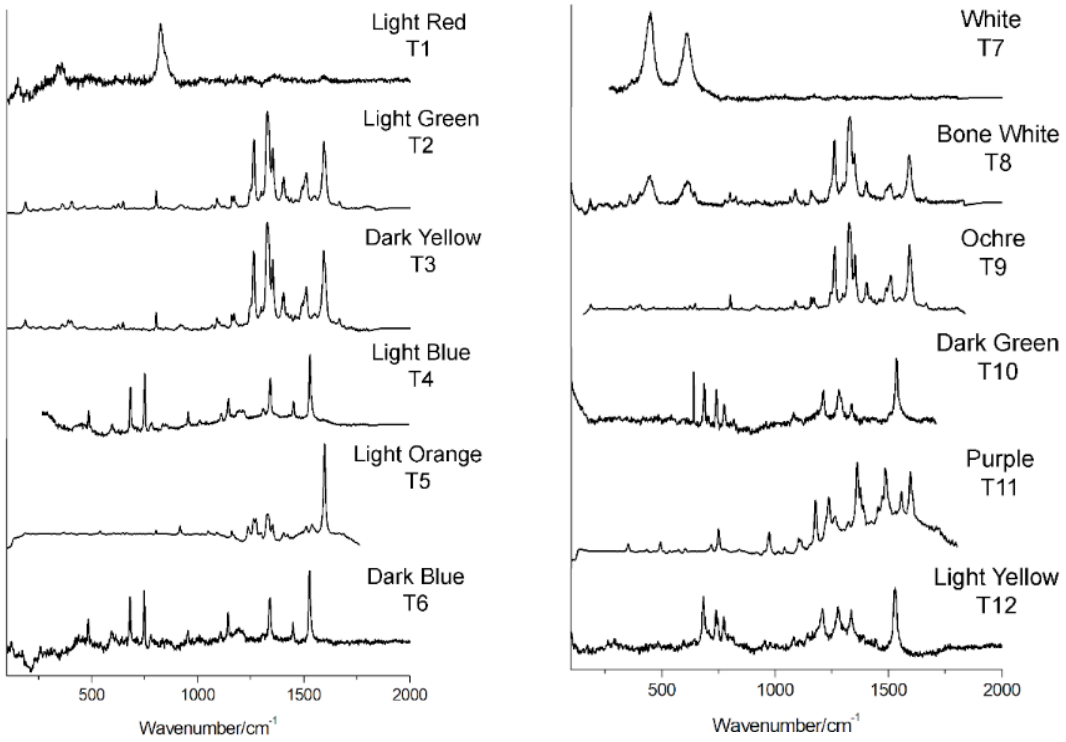


Figure 5: Raman spectra of the sampled inks in the main hall of the Punta Begoña Galleries, with T1 being the most external layer, and T12 the innermost layer

Table 4
Compounds identified by means of Raman spectroscopy in the main hall samples

Name	Composition	Raman Bands
(T1) Light Red	Lead molybdate chromate (PR104)	151, 337, 357, 823s, 849sh cm^{-1}
(T2) Light Green	Acetoacetic arylide (PY74)	187, 224, 260, 319, 361, 404, 465, 524, 601, 624, 646, 802, 846, 917, 1018, 1067, 1089, 1160, 1171, 1264s, 1299, 1327vs, 1353s, 1404, 1423, 1439, 1459, 1512, 1551, 1593s, 1668 cm^{-1}
(T3) Dark Yellow	Acetoacetic arylide (PY74)	186, 260, 359, 402, 464, 524, 601, 623, 646, 742, 802, 846, 917, 1017, 1066, 1089, 1160, 1170, 1263s, 1298, 1327vs, 1352s, 1403, 1423, 1439, 1459, 1490, 1511, 1551, 1593s, 1668 cm^{-1}
	Goethite	301, 387, 549 cm^{-1}
(T4) Light Blue	Phthalocyanine (PB15)	290, 483, 596, 642, 682s, 719, 748vs, 782, 834, 846, 954, 1009, 1110, 1144s, 1197, 1210, 1307, 1342, 1451, 1529vs cm^{-1}
(T5) Light Orange	Disazo pyrazolone (PO34)	294, 369, 392, 539, 669, 768, 915, 1048, 1159, 1188, 1237, 1273, 1289, 1298, 1422, 1475, 1537, 1598vs, 1654 cm^{-1}
	Acetoacetic arylide (PY74)	186, 224, 259, 358, 403, 600, 322, 645, 801, 953, 1018, 1089, 1124, 1159, 1263, 1298, 1327vs, 1351, 1404, 1438, 1498, 1511, 1598s, 1666 cm^{-1}
(T6) Dark Blue	Phthalocyanine (PB15)	290vw, 483, 594, 643vw, 680s, 715vw, 748s, 780, 834, 855, 954, 1008vw, 1108, 1144, 1195, 1307, 1342s, 1450, 1529vs cm^{-1}
(T7) White	Titanium white	448, 606 cm^{-1}
	Alkyd binder	999, 1042, 1172, 1308, 1599 cm^{-1} (very weak)
(T8) Bone White	Acetoacetic arylide (PY74)	186, 264, 314, 359, 404, 645, 802, 827, 915, 953, 1018, 1066, 1090, 1125, 1160, 1263vs, 1300, 1331vs, 1351vs, 1403, 1511, 1549, 1592vs, 1637, 1667 cm^{-1}
	Titanium white	447s, 612s cm^{-1}
	Alkyd binder	1003 cm^{-1} (very weak)
(T9) Ochre	Acetoacetic arylide (PY74)	187, 359, 403, 465, 601, 624, 646, 802, 916, 1018, 1067, 1089, 1160, 1171, 1264s, 1298, 1327vs, 1353s, 1404, 1423, 1439, 1460, 1490, 1512s, 1552, 1593s, 1668 cm^{-1}
	Goethite	385 cm^{-1}
(T10) Dark Green	Phthalocyanine (PG7)	100, 224, 685s, 740s, 774s, 817, 959, 1080, 1212s, 1281s, 1337s, 1535vs cm^{-1}
(T11) Purple	Bona (PR48 Na or PR48:2 Ca)	350, 412, 493, 532, 572, 601, 717, 748, 972, 1040, 1104, 1113, 1177, 1236, 1236, 1263, 1323, 1361, 1375, 1387, 1454, 1473, 1486, 1557, 1596 cm^{-1}
(T12) Light Yellow	Phthalocyanine (PG7)	100, 264, 292, 684, 740, 773, 956, 1081, 1209, 1277,1337, 1442, 1529s cm^{-1}

The study of each ink allowed its identification (Figure 5 and Table 4); starting from the external layer, the T1 ink presented lead in the elemental analysis, and by Raman spectroscopy, lead molybdate chromate pigment (PR104) was found, as can be seen in Figure 5 [15]. This pigment was composed of lead chromate, molybdate and sulphate ($PbCrO_4$-$bMoO_4PbSO_4$) [25]. The next two inks, T2 and T3, were identified as acetoacetic arylide (PY74), also observed in the measured commercial sprays. Even so, the T3 pigment was mixed with iron hydroxide, goethite, as it was assumed in view of the presence of iron in the elemental analysis [16]. This mixture possibly was formulated to obtain the desired tone. Following the layers, the next layer, T4, was identified as copper phthalocyanine (PB15), also identified in the commercial inks, although in this case copper was not found in the elemental analysis [15]. The T5 ink was clearly differentiated, since it was a light orange colour, and was identified as a mixture of two pigments, disazo pyrazolone PO34 and the previously mentioned PY74 [15]. Deeper, the T6 layer, again with a blue colour, showed the PB15 pigment, and in this measurement, copper was identified, as expected [15]. The next three layers (T7-T9) seemed to be very related. In the T7 layer, rutile was identified as a white pigment, along with the alkyd binder observed in the commercial sprays. The next colour was "Bone White" and was considered as a different layer, and in this case titanium oxide, rutile, as white pigment, PY74 as yellow pigment, and a weak band at 1003 cm^{-1}, which could be related to the binder, were identified [15, 16]. The differences between both inks were clear, however; the T9 was an ochre colour, and was identified as PY74 and iron oxide, goethite, confirmed by the major presence of iron in the elemental analysis [15, 16]. So, the T8 layer could be a mixture of the other two layers (T7 and T9). Regarding the elemental analysis, there were some differences, but not enough to confirm if it was a different layer. Continuing, as in the case of commercial sprays, the green colour T10 seemed to be phthalocyanine due to the presence of copper and chloride in the elemental analysis. This observation was confirmed by Raman spectroscopy, which identified PG7 pigment, the copper-based phthalocyanine also observed in the commercial sprays [15]. This pigment also was identified in the final, inner layer, T12. However, the elemental analysis did not help in this case. Finally, the remaining ink, T11, was identified as PR48 or PR48:2 [15]. These pigments are classified as BONA, a name derived from β-hydroxynaphthoic acid, also known as 3-hydroxy-2-naphthoic acid. BONA is the ink component of different diazotized amines containing salt-forming groups [26]. This compound could precipitate with different cations, and depending on them, it could have a secondary number (PR48:X). In this case, the Raman spectra fitted better with the PR48 and PR48:2 pigments (the pigments with sodium and calcium, respectively) [15]. In this sense, the band observed at 1113 cm^{-1} was not present in the literature, but as in the case of the commercial spray "Cologno N. Green RV-237", it seemed to be the result of a better resolution obtained in this study, allowing us to identify this band, which can be found in the literature as a shoulder of the 1177 cm^{-1} Raman band. Regarding the elementary analysis, calcium was present, but it was not possible to avoid its presence as the effect of external agents or even as part of substrate, since it was the inner layer, and sodium is not well observed by XRF, making it impossible to determine exactly which pigment it was. Nevertheless, it did not have a negative impact on the restoration process.

In summary, the composition of the sprays was not just organic, but a mixture of organic and inorganic compounds. In addition, there were some pigments, such as PY74 and PB15, that seemed to be the most common [15]. For restoration purposes, and considering the number of ink layers in the main hall, one or more products will most likely be needed to remove them.

4. Conclusions

The existing databases have been slightly expanded, including elementary information, missed Raman bands in the literature, and the commercial brands. However, more efforts should be made in this regard, above all to ensure that all this information is provided for free through open-access publications. For this reason, the creation of databas-

es of specific sprays, not only of pure pigments, is recommended to help understand them correctly. These databases could help in restoration works of street art, and even to determine the authorships, or in restoration of vandalism actions despite the trademark's secrecy.

Regarding the specific results of this work, even if the characterization of the sampled inks was complicated, it allowed us to identify most of the components of the sprays studied, not just the pigments, despite not having much information from the manufacturers. This information will serve to direct the graffiti-removal protocols to use the most appropriate products. Moreover, the relevance of the elemental analysis, which is generally avoided in graffiti analysis, provided relevant information, and assisted in the identification of the used pigments and other compounds, and also in observing the impact suffered by the building under study.

Thus, this work points out the importance of the study of modern inks before their removal and the need to promote open-access databases that include the elemental and brand information to help in future characterization and diagnostic practices.

5. Supplementary Materials

The following are available online at www.mdpi.com/xxx/s1, Table S1: A listing of all the .spc spectral data.

6. Author Contributions

Experimental analysis and interpretation, I.E., N.P.-T. and E.L.; writing-original draft, N.P.-T.; supervision, M.D.R.-L., G.A., and J.M.M.; project administration, G.A. and J.M.M.; funding acquisition, G.A. and J.M.M. All authors have read and agreed to the published version of the manuscript.

7. Funding

This work was supported by the DEMORA (Grant No. PID2020-113391GB-I00) project funded by the Spanish Agency for Research AEI (MICINN/FEDER-UE), and by the cooperation agreement between the University of the Basque Country and the City Council of Getxo (OTRI 2019-0316).

8. Acknowledgments

The authors wish to acknowledge the professional support of the Interdisciplinary Thematic Platform from CSIC Open Heritage: Research and Society (PTI-PAIS).

9. References

1. Lama, E.; Prieto-Taboada, N.; Etxebarria, I.; Bermejo, J.; Castro, K.; Arana, G.; Rodrí-guez-Laso, M.D.; Madariaga, J.M. Spectroscopic characterization of xx century mural paintings of punta begoña's galleries under conservation works. *Microchem. J.* **2021**, *168*, 106423.
2. Galerías Punta Begoña. Galerías Punta Begoña. 2019. Available online: https://puntabegona.getxo.eus/ (accessed 17/07/2021).
3. European Commission. Real Long-Term Working Conditions of Anti-Graffiti and Self-Cleaning Coatings for Their Imple-Mentation in the Protection of the European Cultural Heritage. 2017. Available online: https://cordis.europa.eu/project/id/622417/es (accessed 17/07/2021).
4. European Commission. Development of a New Antigraffiti System, Based on Traditional Concepts, Preventing Damage of Architectural Heritage Materials. 2012. Available online: https://cordis.europa.eu/project/id/513718/reporting/es (accessed 17/07/2021).
5. European Commission. InnovativE Anti-Graffiti Product for Application in the Cultural Heritage of Europe. 2012. Available online: https://cordis.europa.eu/project/id/513718/reporting (accessed 17/07/2021).
6. Gomes, V.; Dionísio, A.; Pozo-Antonio, J.S. Conservation strategies against graffiti vandalism on Cultural Heritage stones: Protective coatings and cleaning methods. *Prog. Org. Coatings* **2017**, *113*, 90-109.
7. Rivas, T.; Pozo, S.; Fiorucci, M.P.; López, A.J.; Ramil, A. Nd:YVO4 laser removal of graffiti from granite. Influence of paint and rock properties on cleaning efficacy. *Appl. Surf. Sci.* **2012**, *263*, 563-572.
8. Cortea, I.M.; Ratoiu, L.; Rădvan, R. Characterization of spray paints used in street art graffiti by a non-destructive multi-analytical approach. *Color Res. Appl.* **2021**, *46*, 183-194.
9. Germinario, G.; van der Werf, I.D.; Sabbatini, L. Chemical characterisation of spray paints by a multi-analytical (Py/GC-MS, FTIR, μ-Raman) approach. *Microchem. J.* **2016**, *124*, 929-939.
10. Sánchez Pons, M. Acercamiento a la evolución histórica y tecnológica de los materiales pictóricos empleados en el grafiti y arte urbano. *GEC* **2016**, *10*, 146-159.

11. Bosi, A.; Ciccola, A.; Serafini, I.; Guiso, M.; Ripanti, F.; Postorino, P.; Curini, R. Bianco, A. Street art graffiti: Dis-covering their composition and alteration by FTIR and micro-Raman spectroscopy. *Spectrochim. Acta A* **2020**, *225*, 117474.
12. Caggiani, M.C.; Cosentino, A.; Mangone, A. Pigments Checker version 3.0, a handy set for conservation scientists: A free online Raman spectra database. *Microchem. J.* **2016**, *129*, 123-132.
13. Madariaga, J.M. Analytical chemistry in the field of cultural heritage. *Anal. Methods* **2015**, *7*, 4848-4876.
14. Madariaga, J.M. Analytical Strategies for Cultural Heritage Materials and their Degradation. *R. Soc. Chem.* **2021**, P001.
15. Fremout, W.; Saverwyns, S. Identification of synthetic organic pigments: the role of a comprehensive digital Raman spectral library. *J. Raman Spectrosc.* **2012**, *43*, 1536-1544.
16. Infrared & Raman Users Group (IRUG). Spectral Database: Organic Dyes and Pigments. 2021. Available online: http://www.irug.org/search-spectral-database/spectra-index?sortHeader=data_type_raman (accessed 17/07/2021).
17. Noguera-Cámara, J.; Amor-García, R.L. *Conservation Issues in Modern and Contemporary Murals*; Sánchez, M., Shank, W., Fuster, L., Eds.; Cambridge Scholars Publishing: UK, 2015; pp. 166-183.
18. Madariaga, I.; Lama, E.; Calparsoro, E.; Prieto-Taboada, N.; Arana, G.; Laso, M.D.R.; Madariaga, J.M. Enhancement and recovery of the tiles affected by atmospheric pollutants in the Galleries of Punta Begoña, Getxo (Bizkaia). *Boletín Soc. Española Cerámica Vidr.* **2019**, *58*, 161-170.
19. Dahlen, M.A. The Phthalocyanines A New Class of Synthetic Pigments and Dyes. *Ind. Eng. Chem.* **1939**, *31*, 839-847.
20. Scherrer, N.C.; Stefan, Z.; Francoise, D.; Annette, F.; Renate, K. Synthetic organic pigments of the 20th and 21st century relevant to artist's paints: Raman spectra reference collection. *Spectrochim. Acta A* **2009**, *73*, 505-524.
21. Wiesinger, R.; Pagnin, L.; Anghelone, M.; Moretto, L.M.; Orsega, E.F.; Schreiner, M.; Pigment and Binder Concen-trations in Modern Paint Samples Determined by IR and Raman Spectroscopy. *Angew. Chem. Int. Ed.* **2018**, *57*, 7401-7407.
22. Honzíček, J.; Matušková, E.; Voneš, Š.; Vinklárek, J. MaterialsHelmet Phthalocyaninato Iron Complex as a Primary Drier for Alkyd Paints. *Materials* **2021**, *14*, 1-15.
23. García-Florentino, C.; Maguregui, M.; Morillas, H.; Balziskueta, U.; Azcarate, A.; Arana, G.; Madariaga, J.M. Port-able and Raman imaging usefulness to detect decaying on mortars from Punta Begoña Galleries (Getxo, North of Spain). *J. Raman Spectrosc.* **2016**, *47*, 1458-1466.
24. Prieto-Taboada, N.; Ibarrondo, I.; Gómez-Laserna, O.; Martinez-Arkarazo, I.; Olazabal, M.A.; Madariaga, J.M. Buildings as repositories of hazardous pollutants of anthropogenic origin. *J. Hazard. Mater.* **2013**, *248-249*, 451-460.
25. Buzzini, P.; Massonnet, G.; Sermier, F.M. The micro Raman analysis of paint evidence in criminalistics: case studies. *J. Raman Spectrosc.* **2006**, *37*, 922-931.
26. Jaffe, E. E., Pigments, Organic. Kirk-Othmer Encyclopedia of Chemical Technology 2004.

iStock

Azukrearen gehiegizko kontsumoa: ondorio larriak eta globalak dituen arazoa

Azukrearen gehiegizko kontsumoak osasun fisikoan zein mentalean duen eragina xehetasun handiz aztertu da, eta ondorioak argiak dira: gaitz koronario gero eta ohikoagoak, depresioa, hiperaktibitatea, alergiak... Ikerketak emandako datuek arazoari aurre egiteko politika eraginkorrak garatzen lagundu dezakete.

Azukrearen kontsumoak osasunean duen eraginari buruzko ikerketa zientifiko nagusiak berrikusi dituzte lanaren egileek, eta datu kezkagarriak azaleratu dira: azukre asko kontsumitzeak, bereziki fruktosa duena, arazo endokrinoak, metabolikoak eta kardiobaskularrak eragiten ditu, nagusiki. Ondorioak ezagunak dira: obesitatea eta gaixotasun koronarioak, besteak beste. Baina bestelako osasun-arazoekin ere erlazionatuta egon daitezkeela ikusi dute: minbizia, gaitz neuropsikiatrikoak, hezurretako kaltzioa galtzea, gibeleko gaitzak eta alergiak.

Minbiziari dagokionez, ikerketatik ondorioztatu da lotuta egon daitekeela bularreko, prostatako, gibeleko, maskuriko eta endometrioko minbiziaren garapenarekin eta hedapenarekin, nahiz eta horri buruzko ikerketa gehiago egin behar den.

«Prebentzioa sustatzeko hezkuntza eta osasun publikoko politika eraginkorrak behar dira»

Gainera, azterlanak gehiegizko azukre-kontsumoaren eta gaitz neuropsikiatrikoen arteko loturari erreparatu dio, eta ondorioztatu du zerikusi zuzena izan dezakeela depresioarekin, arretaren defizitagatiko nahasmenduarekin eta hiperaktibitatearekin. Ikertzaileek dietaren eta osasun mentalaren artean dagoen erlazio zuzena azpimarratu dute, eta horrek esku-hartze sozialerako bide berriak ireki ditzakeela uste dute.

Haur eta gazteak jomugan

Ondorioak ikusita, ikerketaren egileek gomendatzen dute azukre gehigarriaren kontsumoa eguneko 25 gramo baino gutxiagora murriztea, eta arreta berezia jarri dute haur eta gazteek kontsumitzen duten edari azukredunen kantitatean. Edari azukredunak, gehienez, astean behin hartzea aholkatu dute, osasunean dituzten ondorio kaltegarriak arintzeko.

Azterketak azukrearen kontsumoari lotutako bizi-ohiturei buruzko datuak ere ematen ditu: edari azukredunak maiz kontsumitzen dituzten pertsonek, askotan, eredu dietetiko desorekatua izaten dute, tabakismorako eta alkohol-kontsumorako joera handiagoa, eta ohitura sedentarioak.

Ebidentzia zientifiko hauek garrantzi kliniko eta sozial handia dute, eta azukre gehiegi kontsumitzearen aurkako prebentzio-estrategiak garatzen lagundu dezakete. Ikerketaren egileek prebentzioaren alde egiten dute, eta, horretarako, hezkuntzaren bidez eragiteko eta osasun publikoko politika eraginkorrak ezartzeko eskatzen dute. Besteak beste, azukrearen gaineko zergak, elikagaien etiketen lege zorrotzak eta publizitateari eta marketinari mugak jartzen dizkioten arauak.

InpAkta 00 (2023) 74-98
https://doi.org/10.1136/bmj-2022-071609

Dietary sugar consumption and health: umbrella review

Yin Huang[1], Zeyu Chen[1], Bo Chen[1], Jinze Li[1], Xiang Yuan[2], Jin Li[1], Wen Wang[3], Tingting Dai[4], Hongying Chen[5], Yan Wang[5], Ruyi Wang[1], Puze Wang[1], Jianbing Guo[1], Qiang Dong[1], Chengfei Liu[6], Qiang Wei[1], Dehong Cao[1] and Liangren Liu[1]

[1] Department of Urology/Institute of Urology, West China Hospital, Sichuan University; [2] Department of Oncology, West China Hospital, Sichuan University; [3] Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University; [4] Department of Clinical Nutrition, West China Hospital, Sichuan University; [5] Research Core Facility, West China Hospital, Sichuan University; [6] Department of Urologic Surgery, UC Davis School of Medicine

ABSTRACT: Objective: To evaluate the quality of evidence, potential biases, and validity of all available studies on di- etary sugar consumption and health outcomes.
Design: Umbrella review of existing meta-analyses.
Data sources: PubMed, Embase, Web of Science, Cochrane Database of Systematic Reviews, and hand searching of reference lists.
Inclusion criteria: Systematic reviews and meta-analyses of randomised controlled trials, cohort studies, case- control studies, or cross sectional studies that evaluated the effect of dietary sugar consump- tion on any health outcomes in humans free from acute or chronic diseases.
Results: The search identified 73 meta-analyses and 83 health outcomes from 8601 unique articles, including 74 unique outcomes in meta-analyses of observational studies and nine unique out- comes in meta-analyses of randomised controlled trials. Significant harmful associations be- tween dietary sugar consumption and 18 endocrine/metabolic outcomes, 10 cardiovascular outcomes, seven cancer outcomes, and 10 other outcomes (neuropsychiatric, dental, hepatic, osteal, and allergic) were detected. Moderate quality evidence suggested that the highest ver- sus lowest dietary sugar consumption was associated with increased body weight (sugar sweetened beverages) (class IV evidence) and ectopic fatty accumulation (added sugars) (class IV evidence). Low quality evidence indicated that each serving/week increment of sugar sweetened beverage consumption was associated with a 4% higher risk of gout (class III evidence) and each 250 mL/day increment of sugar sweetened beverage consumption was associated with a 17% and 4% higher risk of coronary heart disease (class II evidence) and all cause mortality (class III evidence), respectively. In addition, low quality evidence sug- gested that every 25 g/day increment of fructose consumption was associated with a 22% higher risk of pancreatic cancer (class III evidence).
Conclusions: High dietary sugar consumption is generally more harmful than beneficial for health, espe- cially in cardiometabolic disease. Reducing the consumption of free sugars or added sugars to below 25 g/day (approximately 6 teaspoons/day) and limiting the consumption of sugar sweetened beverages to less than one serving/week (approximately 200-355 mL/week) are recommended to reduce the adverse effect of sugars on health.
Systematic review registration: PROSPERO CRD42022300982.

* **Corresponding author / Harremanetan jartzeko:** Liangren Liu. Department of Urology/Institute of Urology, West China Hospital, Sichuan University, Chengdu, China. e-mail: liuliangren@scu.edu.cn. https://orcid.org/0000-0001-7597-6241 **Contributed by** Idoia Labayen **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Liu, Liangren *et al.* «Dietary sugar consumption and health: umbrella review», *BMJ* | (2023) 381:e071609 (https://doi.org/10.1136/bmj-2022-071609)

1. Introduction

As an important component of the human diet, sugars have been shown to be harmfully associated with a variety of risk factors for decades, mainly including obesity,[1 2 3] diabetes,[4 5 6] cardiovascular disease,[7 8 9 10] hyperuricaemia,[11] gout,[11 12 13] ectopic fatty accumulation,[14 15] 16 dental caries,[17] and some cancers.[18 19 20 21] According to the latest report of the World Health Organization and the Food and Agriculture Organization of the United Nations, sugars include monosaccharides, disaccharides, polyols, and free sugars, of which free sugars are identified as all monosaccharides and disaccharides added to foods by the manufacturer, cook, or consumer and sugars naturally present in honey, syrups, and fruit juices.[3 22] In addition, another important group of sugars, added sugars, has been proposed in the Dietary Guidelines for Americans and has been defined as all monosaccharides and disaccharides used in processed and prepared foods and drinks and sugars added to foods but not naturally occurring sugars such as in fruits and fruit juices (table 1).[23]

In recent years, many studies have focused on the adverse effects of sugar sweetened beverages on human health, given the substantial contribution of these drinks to total added sugar or free sugar intake and the rapidly increasing rate of their consumption.[24 25 26] Generally, sugar sweetened beverages are the largest source of added sugars, including carbonated and noncarbonated soft drinks, fruit drinks, and sports and energy drinks.[27] Previous surveys have shown that consumption of sugar sweetened beverages is declining in many developed countries, although consumption levels remain high.[27 28] However, the consumption of sugar sweetened beverages is still increasing in many developing countries, which may be attributed to their increased availability accompanied by economic development.[29] The 2007 annual report of the Coca-Cola company revealed that the consumption of sugar sweetened beverages in India and China increased by 14% and 18%, respectively, in one year.[30] In 2018 a cross sectional survey conducted among Chinese primary and junior high school students showed that sugar sweetened beverages provide 10-15% of the total calorie consumption of school students.[31] Data from the National Health and Nutrition Examination Survey (NHANES) showed that, in 2009-10, sugar sweetened beverage consumption contributed 8% and 6.9% of daily energy intake among children/adolescents and adults, respectively, in the US.[32] Additionally, a global survey conducted in 2010 reported that a total of 180 000 adiposity associated deaths could be attributed to the consumption of sugar sweetened beverages around the world.[33] All of these findings promote the development of policies worldwide to limit sugar consumption, including sugars taxes, food labelling laws, and restrictions on advertising and marketing.[34 35 36 37] Meanwhile, national and international organisations such as WHO, the US Department of Agriculture, and the US Department of Health and Human Services have recommended reducing the consumption of free sugars or added sugars to less than 10% of total daily energy intake.[23 38]

Although many meta-analyses of observational studies and randomised controlled trials focused on the associations between sugar consumption and a range of health outcomes have been published in recent decades, deficiencies in the study design, varying measurements of dietary sugar consumption, inconsistent findings, and different definitions of exposure make drawing definitive conclusions difficult. Therefore, before developing detailed policies for sugar restriction, the quality of existing evidence on the associations of dietary sugar consumption with all health outcomes needs to be comprehensively evaluated. To evaluate the quality of evidence, potential biases, and validity of all studies available on dietary sugar intake and any health outcomes, we did an umbrella review of meta-analyses on this topic.

2. Methods

2.1. Umbrella review methods

We systematically searched, extracted, and analysed large amounts of data from published sys-

tematic reviews and meta-analyses that research the associations between various health outcomes and dietary sugar consumption.[39] [40] Generally, dietary sugar consumption could be measured through the specific proportions of sugars in foods or a percentage of total energy and combined in meta-analyses.[3] Therefore, we excluded simple systematic reviews without meta-analyses from our umbrella review. We prospectively registered this umbrella review in PROSPERO (CRD42022300982) (https://www.crd.york.ac.uk/PROSPERO/).

2.2. *Literature search*

We searched PubMed, Embase, Web of Science, and the Cochrane Database of Systematic Reviews from inception through January 2022 (last update) for systematic reviews and meta-analyses of randomised controlled trials and observational studies. We searched the data-bases through a combination of Medical Subject Headings terms, keywords, and variations of text words associated with sugars following the Scottish Intercollegiate Guidelines Network's guidance for literature searching: (sugars OR sugar) AND (systematic review OR meta-analysis).[41] Two authors (YH and ZYC) separately conducted electronic searches to screen the titles and abstracts retrieved from the databases and identified meta-analyses that met the inclusion criteria by full text reading. Any discrepancy in the literature screening between the two reviewers was resolved by a third author (LRL). We hand searched meta-analyses and reviews from the reference lists of all included articles to identify studies that might have been missed.

2.3. *Eligibility criteria*

We identified dietary sugar consumption as the intake of total sugars and the consumption of a component of total sugars (monosaccharides, disaccharides, polyols, free sugars, or added sugars), which are expressed in absolute amounts or as a percentage of total energy, or the intake of sugar sweetened beverages or foods (table 1).[3] We included systematic reviews and meta-analyses of randomised controlled trials, cohort studies, case-control studies, or cross sectional studies that evaluated dietary sugar consumption in humans free from acute or chronic diseases. Meta-analyses were eligible for inclusion when they compared the effects of different dietary sugar consumption on the same health outcome through relative risks, odds ratios, hazard ratios, weighted mean differences, or standardised mean differences. We included meta-analyses when the exposure was total sugars, monosaccharides, disaccharides, polyols, free sugars, added sugars, or sugar sweetened beverages or foods. We extracted data on individual outcomes separately if two or more health outcomes were reported in a study. If more than one study published more than 24 months apart was conducted on the same dietary sugar exposure and health outcomes, we included the most recent study for data extraction, which is generally the study with the largest sample size. If more than one study was conducted within the same 24 month period, we included the meta-analysis with the largest number of prospective cohort studies and randomised controlled trials (a study with a higher AMSTAR score was included if the number of prospective studies was equal).[42] [43] Furthermore, if the most recent study did not do dose-response analysis, whereas another study did, we included both studies for data extraction.

The exclusion criteria for these umbrella reviews included meta-analyses of the association between carbohydrates, non-nutritive sweeteners, and artificially sweetened beverages and health outcomes; meta-analyses evaluating the therapeutic or metabolic effects of short term sugar supplementation; meta-analyses that evaluated the effects of dietary sugar consumption on health outcomes in certain disease populations; randomised controlled trials that aimed to achieve isoenergetic replacement of sugars with other forms of carbohydrate; studies with insufficient data to evaluate sugar consumption from sugar containing foods (such as honey, apples, chocolate, ice cream, 100% fruit juice); and non-English studies and animal and cell culture studies.

2.4. *Data extraction*

Two reviewers (YH and ZYC) independently extracted the following information from each el- igible

study: first author's name, publication year, type of dietary sugar consumption (total sugars, monosaccharides, disaccharides, polyols, free sugars, added sugars, sugar sweetened beverages or foods), measurement of dietary sugar consumption, health outcome, number of included studies, number of cases and total participants, study design (cross sectional, case-control, cohort, and randomised controlled trial), comparison (high versus low, never/low versus moderate/high, any versus none, or extra increment of sugars per day (or week) versus none), and estimated summary effect (risk ratio, odds ratio, weighted mean difference, and standardised mean difference with 95% confidence intervals). Furthermore, we extracted the model of effect (random and fixed), heterogeneity (I^2 statistic and Cochran's Q test P value), and publication bias assessment (P value of Egger's test or funnel plot). If dose-response analysis and subgroup analysis were conducted, we also extracted the non-linearity tests' P value and results of subgroup analysis in meta-analyses. If a meta-analysis was conducted on both cohort and case-control/cross sectional studies and stratification analysis was conducted through study design, we selected the cohort design subanalysis results for data extraction or reanalysed. Any disagreement was determined by a third author (LRL).

2.5. *Quality assessment of methods and evidence*

Two reviewers (YH and ZYC) evaluated the methodological quality of the included articles by using AMSTAR (a measurement tool to assess systematic reviews), a valid and dependable measurement tool in assessing the quality of systematic reviews and meta-analyses.[42 44] In addition, according to the Grading of Recommendations, Assessment, Development and Evaluation (GRADE), we evaluated evidence of each health outcome and graded it as "high," "moderate," "low," or "very low" quality to draw conclusions.[45] Additionally, we classified evidence of outcomes into four categories following the evidence classification criteria: class I (convincing evidence), class II (highly suggestive evidence), class III (suggestive evidence), class IV (weak evidence), and NS (non-significant).[46 47 48] Table 2 shows detailed criteria of evidence classification.

2.6. *Data analysis*

We reanalysed the risk ratio, odds ratio, weighted mean difference, or standardised mean difference with 95% confidence intervals through random or fixed effects models and calculated the I^2 statistic, P value of Cochran's Q test for heterogeneity, and P value of Egger's regression test (at least 10 studies were included) for small study effects in each included meta-analysis that reported the metric, number of cases, and participants of the included original studies.[49 50 51] For outcomes classified as class I or II, we did sensitivity analysis if sufficient data were available to assess whether the credibility of the evidence varied when some component studies were excluded. We also extracted dose-response associations between dietary sugar consumption and various health outcomes from the included meta-analyses, if available.

Moreover, if the latest meta-analysis did not include the original articles that were included by other meta-analyses, we combined the data of these meta-analyses and did a reanalysis. We assessed agreement statistics between two authors (YH and ZYC) regarding study selection by using Cohen's κ statistics and associated 95% confidence interval. We interpreted magnitude of agreement by following guidelines reported by Landis and Koch: slight (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and almost perfect agreement (0.81-1.00).[52] In addition, if a meta-analysis reported the estimated effect by combining observational studies with randomised controlled trials, we reanalysed the estimated effects for observational studies and randomised controlled trials separately. If we could not do a reanalysis from a meta-analysis, we extracted summary data and assessed heterogeneity and publication bias from the meta-analysis as far as possible. We identified a P value <0.10 as statistically significant for heterogeneity tests. For other tests, we considered a P value <0.05 to be significant. We used Review Manager version 5.3 for evidence synthesis, Stata version 12.1 for Egger's test and sensitivity analysis, and IBM SPSS Statistics version 25 for Cohen's κ statistics.

2.7. *Patient and public involvement*

Patients and the public were not involved in the planning, design, and implementation of the study, as this study used secondary data. No patients were asked to advise on interpretation or writing up of the manuscript.

3. Results

3.1. *Characteristics of meta-analyses*

Figure 1 shows the flowchart of the literature search and selection process. After a systematic literature search, we identified 8601 unique articles. Application of our inclusion criteria yielded total of 73 meta-analyses, including 67 meta-analyses of observational studies and six meta-analyses of randomised controlled trials. Agreement between the two reviewers (YH and ZYC) for study selection was almost perfect (κ = 0.906, 95% confidence interval 0.859 to 0.953; P < 0.001). We extracted 74 unique outcomes in meta-analyses of observational studies and nine unique outcomes in meta-analyses of randomised controlled trials. Meta-analyses of randomised controlled trials included only change in body weight (sugar sweetened beverages), liver fat accumulation, muscle fat accumulation, change in body mass index, change in body weight (fructose), postprandial triglycerides, serum uric acid, intrahepatocellular lipids, and alanine aminotransferase. Figure 2 shows the significant dose-response relations between dietary sugar consumption and multiple health outcomes. The other forest plots show the significant non-dose-response relations between dietary sugar consumption and endocrine/metabolic (fig 3), cardiovascular (fig 4), cancer (fig 5), and other outcomes (fig 6).

The full versions of the associations between dietary sugar consumption and each outcome are shown in supplementary tables A-D.

Most of the included meta-analyses focused on the associations between dietary sugar consumption and endocrine/metabolic diseases (n = 28), followed by cancer (n = 25), cardiovascular diseases (n = 17), neuropsychiatric diseases (n = 3), dental diseases (n = 2), and other diseases (n = 8) (fig 7). Dietary sugar exposure included sugar sweetened beverages (n = 58), fructose (n = 11), sucrose (n = 4), lactose (n = 1), added sugars (n = 4), free sugars (n = 1), and total sugars (n = 4). Significance was reached for 45 harmful associations and four beneficial associations. The remaining 34 outcomes were either harmfully or beneficially associated but did not reach significance. After quality assessment of evidence through GRADE and evidence classification criteria, most of the 83 outcomes were classified as "low" or "very low" quality and III, IV, or NS evidence class. Only four (5%) endocrine/metabolic outcomes were classified as "moderate" quality. Three (4%) endocrine/metabolic outcomes, two (2%) cardiovascular outcomes, and three (4%) other outcomes were graded as class IIB. No "high" quality or class I evidence was found in this umbrella review.

3.2. *Endocrine and metabolic outcomes*

3.2.1. Low and moderate quality evidence

A meta-analysis of six randomised controlled trials found that sugar sweetened beverage consumption was significantly associated with increased body weight for highest versus lowest consumption (weighted mean difference 0.85, 95% confidence interval 0.50 to 1.20) (moderate; IV (the quality of evidence is expressed as "GRADE, evidence class")).[53] In addition, any versus no added sugar consumption was associated with increased liver fat accumulation (standardised mean difference 0.93, 95% confidence interval 0.64 to 1.21) (moderate; IV) and muscle fat accumulation (standardised mean difference 0.63, 0.23 to 1.04) (moderate; IV).[54] Another dose-response meta-analysis showed that a one serving/week increment in artificially sweetened beverages was associated with a 4% higher risk of gout (risk ratio 1.04, 95% confidence interval 1.02 to 1.07) (low; III).[13] Furthermore, comparison of higher sugar sweetened beverage consumption with non-sugar sweetened beverage consumption indicated a 55% (odds ratio 1.55, 95% confidence interval 1.32 to 1.82) increased risk of obesity in children associated with higher consumption (low; II).[3] Sugar sweetened beverage consumption was

also linked with an increased body mass index in children.[53] The authors conducted a dose-response analysis and showed that body mass index in children increased by 0.07 units for every one serving/day increment of sugar sweetened beverages (weighted mean difference 0.07, 0.01 to 0.12) (low; IV).[53] Evidence from this umbrella review suggests that fructose intake was not associated with serum uric acid (moderate; NS)[55] or changes in body weight (low; NS) (fig 2; fig 3).[56]

Very low quality evidence Dose-response analysis based on seven cohort studies showed that a one serving/day increment of sugar sweetened beverages was associated with a 0.22 kg weight gain in one year (weighted mean difference 0.22, 0.09 to 0.34).[53] Furthermore, the risk of gout increased by 35% (risk ratio 1.35, 1.18 to 1.55) for the highest versus lowest sugar sweetened beverage consumption.[11] The highest versus lowest sugar sweetened beverage consumption was also significantly associated with a 35% (risk ratio 1.35, 1.19 to 1.52) higher risk of hyperuricaemia.[11] In addition, another pooled analysis found that participants with the highest sugar sweetened beverage consumption had 0.18 mg/dL greater concentrations of serum uric acid than did those with the lowest consumption (weighted mean difference 0.18, 0.11 to 0.25).[57] Similarly, the highest fructose intake could also increase the risk of gout (risk ratio 1.62, 1.28 to 2.03)[58] and hyperuricaemia (odds ratio 1.85, 1.66 to 2.07)[59] compared with the lowest consumption.

The most recent meta-analysis found a 1.46 mg/dL (weighted mean difference −1.46, −2.25 to −0.67) decrement of high density lipoprotein cholesterol for the highest versus lowest sugar sweetened beverage consumption.[60] Subgroup analysis indicated that the highest versus lowest sugar sweetened beverage consumption was associated with lower high density lipoprotein cholesterol in studies conducted in the US (weighted mean difference −2.85, −4.09 to −1.61) but was associated with higher high density lipoprotein cholesterol in studies conducted in Europe/Oceania (weighted mean difference 1.65, 0.26 to 3.05).[60] The highest versus lowest sugar sweetened beverage consumption was also significantly associated with increased low density lipoprotein cholesterol (weighted mean difference 1.21, 0.23 to 2.20) and decreased total cholesterol (−2.49, −2.89 to −2.10).[60] After stratification by region, no significant association between sugar sweetened beverage consumption and low density lipoprotein cholesterol was detected in the US, Europe/Oceania, and Asia,[60] whereas the highest versus lowest sugar sweetened beverage consumption was associated with lower total cholesterol concentrations in studies conducted in the US/Europe (weighted mean difference −2.47, −2.88 to −2.07) but not in Asia.[60]

The risk of metabolic syndrome was increased by 14% (risk ratio 1.14, 1.05 to 1.23) for a 355 mL/day increment of sugar sweetened beverages, with no evidence for departure from linearity.[61] In addition, a meta-analysis including 56 579 participants and 11 821 incident cases of obesity showed an adverse linear dose-response association between sugar sweetened beverage consumption and the risk of obesity.[1] Each 250 mL/day increment in sugar sweetened beverage consumption was associated with a 12% (risk ratio 1.12, 1.05 to 1.19) higher risk of obesity, and this association also remained after adjustment for energy intake (1.13, 1.09 to 1.18) and physical activity (1.14, 1.05 to 1.25).[1] Moreover, a meta-analysis of 16 cohort studies found that with each one serving/day increment of sugar sweetened beverage consumption, the risk of developing type 2 diabetes mellitus increased by 27% (risk ratio 1.27, 1.15 to 1.41).[6] By contrast, an 8% (risk ratio 0.92, 0.85 to 0.99) lower risk of type 2 diabetes mellitus for each 25 g/day increment in sucrose intake was confirmed in dose-response analysis based on six cohort studies.[62] The highest versus lowest sugar sweetened beverage consumption was also significantly associated with a higher risk of latent autoimmune diabetes in adults (odds ratio 1.26, 1.12 to 1.41) (fig 2; fig 3).[30]

We found no significant association between sugar sweetened beverage consumption and changes in body mass index in adults,[63] triglycerides,[60] or large waist circumference.[64] Fructose intake was not associated with postprandial triglycerides or type 2 diabetes mellitus.[62 65] Total sugar consumption was also not associated with type 2 diabetes mellitus (supplementary table A).[62]

3.3. *Cardiovascular outcomes*

Low quality evidence In a single article,[10] a positive association between sugar sweetened beverage consumption and the risk of coronary heart disease was observed. Dose-response analysis showed that each 250 mL/day increment of sugar sweetened beverage consumption was positively associated with a 17% (risk ratio 1.17, 1.11 to 1.23) higher risk of coronary heart disease (low; II).[10] In addition, extreme category analysis showed that the highest versus lowest sugar sweetened beverage consumption was associated with an increased risk of myocardial infarction (risk ratio 1.19, 1.09 to 1.31) (low; III).[66] Low quality evidence suggests that fructose intake was not associated with the risk of hypertension (low; NS) (fig 2; fig 4).[67]

Very low quality evidence Except for a beneficial association between sucrose intake and cardiovascular disease mortality, all categories of dietary sugar exposure were adversely associated with various cardiovascular outcomes. A recent dose-response meta-analysis showed that each 250 mL/day increment of sugar sweetened beverage consumption was positively associated with a 7% (risk ratio 1.07, 1.02 to 1.12) higher risk of stroke.[10] Another meta-analysis of seven cohort studies with 329 791 participants and 16 999 cases found that each one serving/day increment of sugar sweetened beverage consumption was linearly associated with an 8% (risk ratio 1.08, 1.02 to 1.14) increased risk of cardiovascular disease.[8] For cardiovascular disease mortality, each serving/day increment of sugar sweetened beverage consumption was also linearly associated with a higher risk (hazard ratio 1.08, 1.04 to 1.12).[68] However, subgroup analysis found that the association between sugar sweetened beverage consumption and cardiovascular disease mortality was not statistically significant among participants from Asia.[68] In a separate meta-analysis in children and adolescents,[69] the highest versus lowest sugar sweetened beverage consumption was shown to be associated with a 1.67 mm Hg (weighted mean difference 1.67, 1.02 to 2.32) increase in systolic blood pressure and a 36% (odds ratio 1.36, 1.14 to 1.63) higher risk of hypertension. In adults, the results from pooled analysis of 13 prospective cohort studies indicated a harmful dose-response association between sugar sweetened beverage consumption and incidence of hypertension.[70] The risk of hypertension was increased by 11% (risk ratio 1.11, 1.09 to 1.13) for a 355 mL/day (1 serving/day) increment in sugar sweetened beverage consumption.[70] Moreover, both fructose (risk ratio 1.08, 1.01 to 1.15) and total sugars (risk ratio 1.09, 1.02 to 1.17) were harmfully associated with the risk of cardiovascular disease mortality for highest versus lowest consumption,[71] whereas a beneficial association between sucrose intake and cardiovascular disease mortality was observed (fig 2; fig 4).[71]

We observed no significant association between sugar sweetened beverage consumption and changes in diastolic blood pressure (children and adolescents)[69] or heart failure.[10] We also observed no significant association between sucrose intake or total sugar consumption and the risk of cardiovascular disease.[71] In addition, added sugar consumption was not associated with the risk of cardiovascular disease mortality (supplementary table B).[71]

3.4. *Cancer*

Low quality evidence A dose-response meta-analysis showed that the risk of hepatocellular carcinoma increased by 100% (risk ratio 2.00, 1.33 to 3.03) for the highest sugar sweetened beverage consumption compared with the lowest (low; IV).[18] Additionally, a meta-analysis conducted by Aune and colleagues found that 25 g/day of fructose intake was linearly associated with a 22% higher risk of pancreatic cancer (risk ratio 1.22, 1.08 to 1.37) (low; III).[72] The association between fructose intake and incidence of pancreatic cancer remained significant in the subgroups of studies that adjusted for smoking, body mass index, red and processed meat consumption, and energy intake, whereas the association was diminished in the subgroups of studies that adjusted for alcohol consumption, diabetes status, or physical activity (fig 2; fig 5).[72]

Very low quality evidence A recent meta-analysis of six observational studies showed a higher risk of breast cancer for highest versus lowest sugar sweetened beverage consumption (risk ratio 1.14,

1.01 to 1.30).[19] In a separate meta-analysis, Li and colleagues found that the highest sugar sweetened beverage consumption might increase the risk of breast cancer mortality by 17% (risk ratio 1.17, 1.03 to 1.34) compared with the lowest.[18] Moreover, a meta-analysis of six cohort studies showed that participants with the highest sugar sweetened beverage consumption had a higher risk of prostate cancer than those with the lowest intake (risk ratio 1.17, 1.07 to 1.28). Dose-response analysis did not detect a significant association.[18] However, we observed a protective association between sugar sweetened beverage consumption and glioma in our umbrella review (risk ratio 0.81, 0.66 to 0.99).[18] In addition, a meta-analysis including 20 cohort studies with 5 505 812 participants observed a positive linear dose-response relation between sugar sweetened beverage consumption and overall cancer risk.[18] The risk increased by 4% for every serving/day increment of sugar sweetened beverage consumption (risk ratio 1.04, 1.01 to 1.09).[18] Furthermore, pooled analysis of 10 cohort studies with 1 239 183 participants found that the highest versus lowest sugar sweetened beverage consumption was significantly associated with a higher risk of overall cancer mortality (risk ratio 1.06, 1.00 to 1.12), without a significant dose-response relation.[18] Stratification by region produced a positive association between sugar sweetened beverage consumption and overall cancer mortality in the North American population (odds ratio 1.08, 1.01 to 1.15) but not in Asia (0.99, 0.81 to 1.22) (fig 2; fig 5).[18]

We observed no significant association between sugar sweetened beverage consumption and the risk of biliary track cancer,[18] bladder cancer,[18] colon cancer,[73] colorectal cancer,[18] colorectal cancer mortality,[18] endometrial cancer,[18] oesophageal cancer,[18] gastric cancer,[18] haematological malignancy,[18] kidney cancer,[18] lung cancer mortality,[18] nasopharyngeal carcinoma,[18] pancreatic cancer,[18] and prostate cancer mortality.[18] In addition, added sugar consumption was not associated with the risk of colorectal cancer.[74] We observed no significant association between sucrose intake and pancreatic cancer.[72] Moreover, lactose intake was not associated with the risk of ovarian cancer (supplementary table C).[75]

3.5. Other outcomes

Low quality evidence A recent meta-analysis of 11 cohort studies suggested that an increment in sugar sweetened beverage consumption of 250 mL/day was associated with a 4% (hazard ratio 1.04, 1.02 to 1.06) higher risk of all cause mortality (low; III).[76] Moreover, a harmful association between sugar sweetened beverage consumption and the risk of depression was observed in a meta-analysis of 10 observational studies (risk ratio 1.31, 1.24 to 1.39) (low; II).[77] No significant association was observed between fructose intake and alanine transaminase concentration (low; NS) (fig 2; fig 6).[78]

Very low quality evidence The highest versus lowest sugar sweetened beverage consumption might increase the risk of asthma in children by 26% (odds ratio 1.26, 1.07 to 1.48).[79] In a single article,[80] both sugar sweetened beverage consumption (odds ratio 1.80, 1.23 to 2.63) and total sugar consumption (1.22, 1.04 to 1.42) were associated with an increased risk of attention deficit/hyperactivity disorder. In addition, the results from a meta-analysis of 10 observational studies showed a significant inverse association between sugar sweetened beverage consumption and bone mineral density in adults (standardised mean difference −0.66, −1.01 to −0.31).[81] Subgroup analysis according to sex showed a significant harmful effect of sugar sweetened beverage consumption on bone mineral density in females (standardised mean difference −0.50, −0.87 to −0.13) but no association in males.[81] For dental diseases, a single article found a harmful association between sugar sweetened beverage consumption and the incidence of dental caries (odds ratio 1.72, 1.41 to 2.09) and dental erosion (1.77, 1.28 to 2.43) when comparing never/low with moderate/high consumption.[17] Additionally, sugar sweetened beverage consumption was positively associated with the risk of non- alcoholic fatty liver disease (risk ratio 1.39, 1.29 to 1.50).[16] Fructose intake was associated with increased intrahepatocellular lipids (standardised mean difference 0.45, 0.18 to 0.72) (fig 2; fig 6).[78]

Sugar sweetened beverage consumption was not associated with the risk of chronic kidney disease.[82]

In addition, maternal increased free sugar intake during pregnancy was not associated with the risk of asthma in offspring (supplementary table D).[83]

3.6. *Heterogeneity*

We reanalysed the heterogeneity in 69% of all health outcomes by a random or fixed effects model. Reanalysis found that approximately 46% of the health outcomes that we reanalysed had significant heterogeneity (I^2 > 50% or P value of Cochran's Q test < 0.1). The heterogeneity of most outcomes could be explained by some potential factors, including setting, region, ethnicity, sex, age, study quality, study design, sample size, duration of follow-up, and adjustment for confounding factors. Of the 26 outcomes that we could not reanalyse, approximately 54% had significant heterogeneity and 4% did not report the results of the heterogeneity evaluation.

3.7. *Assessment of risk of bias*

We conducted Egger's test for 23% of the outcomes in our reanalysis, which found evidence of publication bias in three outcomes —type 2 diabetes mellitus (sugar sweetened beverages) (P = 0.016), overall cancer risk (P = 0.005), and hypertension in adults (sugar sweetened beverages) (P = 0.02). For outcomes that we could not reanalyse, publication bias was detected for cardiovascular disease mortality (sugar sweetened beverages), non-alcoholic fatty liver disease, obesity in adults, and change in body weight (one year) by statistical test or funnel plot. The remaining outcomes did not show significant publication bias or did not report an evaluation for publication bias.

3.8. *AMSTAR, GRADE, and evidence classification*

The median AMSTAR score of all health outcomes was 8 (range 3-11; interquartile range 8-9.25) (supplementary table E). Supplementary table F provides the detailed AMSTAR scores for each outcome. All evidence from meta-analyses of cohorts, population based case-control studies, and cross sectional studies is graded as "low" or "very low" quality owing to the observational study design and factors for quality downgrade (significant risk of bias, inconsistency, indirectness, imprecision, and potential publication bias). Among the nine meta-analyses of randomised controlled trials, four (liver fat accumulation, muscle fat accumulation, serum uric acid (fructose), and change in body weight (sugar sweetened beverages)) were downgraded as "moderate" quality given the imprecision, and the remaining (alanine transaminase, intrahepatocellular lipids, postprandial triglycerides, change in body mass index in adults, and change in body weight (fructose)) were downgraded as "low" or "very low" owing to the risk of bias, inconsistency, indirectness, or imprecision (supplementary table E). Supplementary Table G shows the detailed GRADE classification for each outcome. In terms of evidence classification, type 2 diabetes mellitus (sugar sweetened beverages), hyperuricaemia (fructose), obesity in children (sugar sweetened beverages), coronary heart disease, hypertension in adults (sugar sweetened beverages), dental caries, depression, and non-alcoholic fatty liver disease were graded as class II. For the remaining 75 outcomes, 15 (18.1%) were graded as class III, 26 (31.3%) were graded as class IV, and 34 (41.0%) were identified as non-significant (supplementary table E). Sensitivity analyses of each outcome graded as class II did not alter the direction or significance of the association.

4. Discussion

4.1. *Principal findings and possible explanations*

Dietary sugar consumption is harmfully associated with multiple health outcomes across various measurements of exposure, including high versus low, never/low versus moderate/high, any versus none, or an extra increment of sugars per day (or week) versus none. We identified 73 meta-analyses and 83 health outcomes from 8601 unique articles, including 74 unique outcomes in meta-analyses of observational studies and nine unique outcomes in meta-analyses of randomised controlled trials.

Dietary sugar consumption had harmful associations with endocrine and metabolic outcomes, in-

cluding changes in body mass index in children,[53] changes in body weight,[53] changes in body weight (one year),[53] gout,[11 13 58] high density lipoprotein cholesterol,[60] hyperuricaemia,[11 59] latent autoimmune diabetes in adults,[30] low density lipoprotein cholesterol,[60] metabolic syndrome,[61] obesity in children,[3] obesity in adults,[1] serum uric acid,[57] type 2 diabetes mellitus,[6] liver fat accumulation,[54] and muscle fat accumulation.[54] In addition, harmful associations between dietary sugar consumption and cardiovascular outcomes were also observed, including coronary heart disease,[10] cardiovascular disease,[8] cardiovascular disease mortality,[68 71] hypertension in children and adolescents,[69] hypertension in adults,[70] myocardial infarction,[66] change in systolic blood pressure in children and adolescents,[69] and stroke.[10] Significant harmful associations between dietary sugar consumption and a higher risk of cancer were observed for breast cancer,[19] breast cancer mortality,[18] hepatocellular carcinoma,[18] prostate cancer,[18] pancreatic cancer,[72] overall cancer risk,[18] and overall cancer mortality.[18] Finally, harmful associations existed between dietary sugar consumption and all cause mortality,[76] asthma in children,[79] attention deficit/hyperactivity disorder,[80] bone mineral density,[81] dental caries,[17] dental erosion,[17] depression,[77] non-alcoholic fatty liver disease,[16] and intrahepatocellular lipids.[78]

In general, no reliable evidence shows beneficial associations between dietary sugar consumption and any health outcomes, apart from glioma (sugar sweetened beverages),[18] total cholesterol (sugar sweetened beverages),[60] type 2 diabetes mellitus (sucrose),[62] and cardiovascular disease mortality (sucrose).[71] However, these favourable associations are not supported by strong evidence, and the interpretation of these results should be done with caution. For the decreased risk of glioma, evidence for this came from only two cohort studies, and no studies have shown that sugar sweetened beverage consumption is a protective factor to lower the incidence of cancer. High sugar sweetened beverage consumption was associated with lower total cholesterol concentrations. However, subgroup analysis indicated that sugar sweetened beverage consumption was associated with higher total cholesterol concentrations in studies with sugar sweetened beverage consumption > 750 g/day and studies involving adolescents. Therefore, potential confounders, including region, sugar sweetened beverage dose, sample size, and sex, should be considered in explaining the association between sugar sweetened beverage consumption and total cholesterol concentrations. In terms of the protective effect of sucrose intake on type 2 diabetes mellitus and cardiovascular disease mortality, we note that sucrose tends to be found more in solid foods than in sugar sweetened beverages, including grains and grain based products, fruit and fruit products, and sweetened dairy and dairy products.[84 85 86] These main sources of sucrose have shown beneficial associations (for example, whole grain cereals, fruit, and yogurt) with type 2 diabetes mellitus and cardiovascular disease mortality.[87 88 89 90 91 92] Therefore, the protective association between sucrose intake and type 2 diabetes mellitus and cardiovascular disease mortality may reflect important contributions from these other food sources rather than sucrose.[62 71] Further large scale, prospective studies are warranted to evaluate the association of sucrose intake with type 2 diabetes mellitus and cardiovascular disease mortality and to clarify the possible underlying mechanisms.

Our umbrella review showed harmful associations between dietary sugar consumption and a range of cardiometabolic diseases, especially weight gain, ectopic fat accumulation, obesity, and cardiovascular disease, which can largely be attributed to excessive consumption of fructose containing sugars. In response to the intake of large carbohydrates, fructose could enhance hepatic lipogenic capacity by inducing hepatic master transcription factors.[93 94 95] Moreover, an animal study found that dietary fructose could be converted to microbial acetate by the gut microbiota, which may enhance hepatic lipogenesis by supplying lipogenic acetyl-CoA independently of ATP citrate lyase.[96] Intermediate products such as diacylglycerols generated during the process of lipogenesis may impair insulin signalling in the liver and peripheral tissues and then lead to insulin resistance.[97] Subsequently, it may promote ectopic fat deposition in the liver and

muscle.[98 99] Dietary fructose may also inhibit fatty acid oxidation in the liver by impairing mitochondrial size and function and acetylation of the rate limiting enzyme.[100] A recent animal study showed that dietary fructose improves the survival of intestinal cells and increases the length of intestinal villus in mouse models, resulting in an expanded surface area of the gut and increased nutrient absorption and adiposity in mice.[101] Furthermore, fructose contained in sugar sweetened beverages is suggested to likely induce the onset of obesity by reducing resting energy expenditure and promoting leptin resistance.[102 103] In addition, sugar sweetened beverages are associated with less satiety compared with solid food containing the same amount of calories, which may stimulate appetite and induce excessive calorie consumption, liver fat accumulation, and insulin resistance in the long term.[104] This hypothesis is confirmed by several clinical trials conducted in healthy adults, which found that sugar sweetened beverage consumption results in more caloric in-take and weight gain than artificially sweetened beverages.[105 106 107] Additionally, a recent double blind, randomised controlled trial carried out in 94 healthy men suggested that consumption of sugar sweetened beverages containing fructose might induce a significant change in the low density lipoprotein particle distribution towards smaller, more atherogenic particles, partially mediating the associations of sugar sweetened beverage consumption with dyslipidaemia and cardiovascular disease.[108]

Another important mechanism to explain the associations between dietary sugar consumption and cardiometabolic diseases involves uric acid synthesis. Many studies have confirmed that excessive fructose consumption can promote uric acid synthesis by inducing degradation ATP to AMP, a substrate for uric acid production.[109 110 111] Fructose phosphorylation in the liver uses ATP to convert fructose into fructose-1-phosphate and leads to phosphate depletion, which limits the regeneration of ATP from ADP. Then, ADP is converted to AMP and consequently induces the synthesis of uric acid.[57] In addition, fructose induced hyperinsulinaemia and insulin resistance may also result in higher serum uric acid by reducing the excretion of uric acid.[110 112 113] Hyperuricaemia is a precursor to gout.[109 110] The positive associations between gout, hyperuricaemia, and other cardiometabolic diseases, such as hypertension, type 2 diabetes mellitus, and cardiovascular disease, have been proposed for a long time.[114 115] Hyperuricaemia has been shown to precede the occurrence of type 2 diabetes mellitus and obesity.[27] Mechanistically, hyperuricaemia could induce renal microvascular alteration, chronic sodium retention, reduction in nitric oxide concentrations in endothelial cells, and the activation of the renin-angiotensin system, which may account for the association between fructose containing sugar consumption and cardiovascular disease.[114 116 117 118]

Until now, the evidence for the association between dietary sugar consumption and the risk of cancer has remained limited and controversial.[27] In 2018 the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) reported that evidence was limited for the associations between consumption of sugars and food containing sugars and the risk of colorectal cancer.[119] However, at the same time, this report recommended reducing or avoiding sugar sweetened beverage consumption for the prevention of breast cancer.[119] Evidence from this umbrella review supports the recommendations from the WCRF/AICR to some extent. In our study, although eight of the 25 cancer outcomes were identified as being positively associated with dietary sugar consumption (seven exposure factors were sugar sweetened beverages, and one was fructose), only evidence of hepatocellular carcinoma (sugar sweetened beverages) and pancreatic cancer (fructose) were rated as "low" quality because of the magnitude of effect or dose-response gradient, and the remaining outcomes were all rated as "very low" quality. As a result, caution is warranted when explaining the significant associations between dietary sugar consumption and some cancer risks.

The effect of dietary sugars on obesity might partly explain their association with the risk of cancer.[21] As mentioned previously, dietary sugar consumption, especially sugar sweetened beverage consumption,

is convincingly associated with the risk of obesity weight gain,[1 3 53] which in turn is regarded as a strong risk factor for various cancers.[21 119] Another pathway mediating the association between dietary sugar consumption and the risk of cancer might involve a high glycaemic index or glycaemic load. The glycaemic index has been associated with the risk of type 2 diabetes mellitus,[120] which may be involved in carcinogenesis of the breast, prostate, liver, bladder, and endometrium.[120 121] Moreover, excessive fructose consumption might lead to intestinal flora disturbance and intestinal barrier deterioration, which promote the development of metabolic endotoxaemia, inflammation, and lipid accumulation, finally leading to colorectal carcinogenesis.[20 122 123] A recent animal study showed that high fructose corn syrup intake could induce intestinal tumourigenesis in mice by expediting glycolysis and de novo lipogenesis. The mice treated with the syrup had a substantially increased tumour size and tumour grade, independent of obesity and metabolic syndrome.[124] Considering the different mechanisms of site specific cancers, further prospective studies that explore the definite associations between sugar consumption and cancer risk for diverse cancer types and ethnic groups are warranted.[27]

On the other hand, dietary sugar consumption has also been shown to be negatively associated with some neuropsychiatric diseases, such as depression and attention deficit/hyperactivity disorder.[77 80] Several biological mechanisms might be involved in these associations.

Data from an animal study showed that a high fructose diet might alter behaviour, hypothalamic-pituitary-adrenal axis function, and the hypothalamic transcriptome in male Wistar rats, inducing anxiety-like behaviour and depressive-like behaviour.[125] Furthermore, sugar consumption has been suggested to stimulate the secretion of endogenous opioids in the nucleus accumbens and to stimulate the dopaminergic reward system.[27] Evidence of sugar dependence in an animal model indicated that similarly to addiction to morphine and cocaine, rats with intermittent sugar intake had decreased concentrations of dopamine D2 receptor mRNA in the nucleus accumbens and showed the characteristics of addictive-like be- haviours called sugar addiction.[27 126]

In addition, the adverse association between sugar consumption and bone mineral density might be attributed to the increased loss of urinary calcium and imbalance in calcium homoeostasis induced by high sugar intake.[127] As well as the negative effect of sugars, phosphate, acidity, and caffeine contained in sugar sweetened beverages are three other major factors that affect bone metabolism.[81] We note that for the link between sugar sweetened beverages and bone mineral density, stratification analysis by gender showed a significant harmful effect of sugar sweetened beverages on bone mineral density in females but not in males.[81] These diverse findings indicated that sugar sweetened beverage consumption had a more detrimental effect on female bone health than on male bone health because women generally have smaller bones and lower bone strength and are therefore more susceptible to osteoporosis.[128] Moreover, the high acidity of sugar sweetened beverages is also thought to be an important factor in promoting dental caries and tooth erosion.[129 130 131]

Of the subgroup analyses conducted in this umbrella review, the most noteworthy is the stratification according to region, as several health outcomes showed a regional discrepancy, including overall cancer mortality, high density lipoprotein cholesterol, low density lipoprotein cholesterol, and total cholesterol. Potential reasons for these discrepancies may include regional differences in sugar consumption and culture. According to the report conducted in 2010 for the quantification of global, regional, and national consumption of sugar sweetened beverages in 187 countries, consumption among Asian countries was lower than that among European and American countries.[33] The consumption of sugar sweetened beverages was highest in the Caribbean and lowest in East Asia and Oceania.[33] In addition, cultural factors have been shown to potentially cause different dietary quality and health inequalities by affecting food preferences or choices.[132] Regional cultural diversity in lifestyle and sociodemographic factors also plays an important role in dietary sugar consumption, which may

partly explain the different relations between sugar consumption and disease risk in ethnically diverse populations.[132] [133] On the other hand, subgroup analyses with adjustment for confounding factors should also be considered. High consumption of sugars, especially sugar sweetened beverages, may be a marker of an unhealthy diet and lifestyle.[9] [66] People who consumed sugar sweetened beverages more frequently were likely to ingest more total and saturated fat, carbohydrate, and sodium and less fruit, fibre, dairy products, and wholegrain foods.[134] [135] [136] [137] [138] This dietary pattern was also associated with more frequent smoking and drinking, lower physical activity levels, and more time spent watching television.[137] [138] Therefore, the role of these confounding factors should be taken into consideration when explaining the association between sugar consumption and burden of disease.

4.2. *Strengths and weaknesses of study and in relation to other studies*

This umbrella review first reported a comprehensive summary of the current evidence from previous meta-analyses of observational studies and randomised controlled trials for the association between dietary sugar consumption and all health outcomes. Given the high levels of dietary sugar consumption worldwide, this study has clinical and social significance for developing preventive strategies against excessive sugar consumption, especially for children and adolescents. This study was carried out on the basis of systematic methods in which independent literature searching, study selection, and data extraction by two authors were involved. If the data were sufficient, we reanalysed the risk ratio, odds ratio, weighted mean difference, or standardised mean difference with 95% confidence intervals through random or fixed effects models and evaluated the heterogeneity and publication bias for each included meta-analysis. Furthermore, we used three standard approaches, AMSTAR, GRADE, and evidence classification criteria, to assess the methodological quality (AMSTAR), strength (GRADE) and classification (evidence classification criteria) of evidence for each health outcome and to evaluate our confidence in the estimates. Interestingly, in our umbrella review, the GRADE rating of several health outcomes was not completely consistent with the results of evidence classification. As we know, evidence classification criteria are a completely objective classification standard, whereas the GRADE rating has a certain degree of subjectivity.[139] Therefore, both the GRADE rating and evidence classification criteria should be considered when evaluating evidence and making recommendations.

Original studies included in meta-analyses used different methods of food intake investigation, including food records, 24 hour dietary recall, food frequency questionnaires, and dietary history. All of these are associated with an unavoidable measurement bias even if validated methods are used.[3] This limitation is common to all major epidemiological studies carried out worldwide in this field.[21] In addition, most studies focused on beverages presweetened before purchase.[9] For instance, in the Nurses' Health Study, coffee with sugars was excluded from sugar sweetened beverages, which might affect the reliability of the association.[137] Similarly, another limitation of our study was that we could not evaluate sugar intake in some foods that potentially contain sugars, such as chocolate and ice cream, because of a failure to extract data on sugar consumption. Furthermore, the types of sugar sweetened beverages and dosage of their consumption varied in the original studies. In this umbrella review, most meta-analyses produced summary effects from original studies that measured exposure to dietary sugars through the number of servings a day. However, in some original studies, the number of millilitres a day, grams a day, times a day, times a week, times a month, servings a week, or servings a month were used to estimate sugar consumption, which may partly explain the origin of heterogeneity in meta-analyses. Therefore, dose-response analysis and stratification analysis by sugar sweetened beverage types were unavailable for most outcomes owing to diverse measurements of sugar sweetened beverage consumption in the original studies. Consumption of sugars in sugar sweetened beverages is generally accompanied by the ingestion of some other chem-

ical compounds, such as 4-methylimidazole,[140 141] pesticides,[142 143] artificial sweeteners,[144] sodium benzoate,[79] and sulfites,[79] which may confuse the effect of sugars and therefore should be regarded as potential confounding factors.

We reviewed details of competing interest and funding disclosures from meta-analyses included in this umbrella review. Only two meta-analyses were funded by companies that produce sugar sweetened beverages.[65 145] Among them, the meta-analysis conducted by Wang and colleagues was selected for data extraction and is shown in summary tables.[65] Therefore, caution is warranted when explaining the non-significant association between fructose intake and postprandial triglycerides. Another meta-analysis was not selected for data extraction,[145] and the list of all meta-analyses not selected for data extraction and reanalysis are available if needed. We did not investigate the original studies included in each meta-analysis and therefore could not confirm whether these studies had a competing interest with companies associated with the sugar industry.[42]

The harmful association between dietary sugar consumption and multiple health outcomes observed in our umbrella review is supported by several large scale prospective cohort stud- ies published in recent years. The first was a large prospective cohort study conducted using the results of the French NutriNet-Santé cohort (2009-17), which included 101 257 participants with an average age of 42.2.[21] During the eight year follow-up period, a total of 2193 cases of cancer were reported, including 693 cases of breast cancer. A harmful association was found between sugar sweetened beverage consumption and the risk of overall cancer (hazard ratio 1.18, 1.10 to 1.27) and breast cancer (1.22, 1.07 to 1.39). No significant association was observed for sugar sweetened beverage consumption and the risk of prostate, colorectal, and lung cancer.[21] In this umbrella review, however, the highest versus lowest sugar sweetened beverage consumption was associated with a 17% increased risk of prostate cancer, without a dose-response gradient. Notably, the non-significant association between sugar sweetened beverage consumption and the risk of colorectal cancer observed both in this study and in our umbrella review was inconsistent with another cohort conducted in women.[20] In the Nurses' Health Study II (1991-2015), the authors prospectively explored the association of sugar sweetened beverage consumption in adulthood and adolescence with the risk of early onset colorectal cancer among 95 464 women. A total of 109 cases of early onset colorectal cancer were confirmed during follow-up. Compared with women who consumed less than one serving a week of sugar sweetened beverages in adulthood, those who consumed two or more servings a day had a 118% higher risk of early onset colorectal cancer (risk ratio 2.18, 1.10 to 4.35). Each one serving a day increment of sugar sweetened beverage consumption was associated with a 16% (risk ratio 1.16, 1.00 to 1.36) increased risk of early onset colorectal cancer.[20] In addition, another prospective cohort study showed that excessive consumption of sugars and sugar sweetened beverage during adolescence was significantly associated with the risk of colorectal adenoma (odds ratio 1.20, 1.04 to 1.39).[146] Each one serving a day increase in sugar sweetened beverage consumption was associated with 11% (odds ratio 1.11, 1.02 to 1.20) and 30% (1.30, 1.08 to 1.55) higher risks of total colorectal adenoma and rectal adenoma, respectively.[146] Given that the association between sugar consumption and colorectal cancer risk remains controversial, further well designed, large scale prospective studies are needed to clarify it.

The positive associations between sugar sweetened beverage consumption and the risk of mortality detected in this umbrella review were supported by a prospective cohort study of 118 363 people followed for 34 years in the US, during which time 36 436 deaths were documented.[147] After adjustment for diet and lifestyle confounders, the consumption of two or more servings of sugar sweetened beverages a day was linked with a 21% (hazard ratio 1.21, 1.13 to 1.28) higher risk of total mortality, a 31% (1.31, 1.15 to 1.50) higher risk of cardiovascular disease mortality, and a 16% (1.16, 1.04 to 1.29) higher risk of cancer mortality.[147] On the other hand, a prospective cohort study of 120 343 UK participants followed for 8.4 years confirmed the harmful association of added sugar consumption with the risk of

type 2 diabetes mellitus.[148] A dietary pattern high in added sugars was associated with a higher incidence of type 2 diabetes mellitus (hazard ratio 1.09, 1.06 to 1.12) after adjustment for confounders.[148] Similar to their findings, we observed a strongly significant association between consumption of sugar sweetened beverages (one of the main sources of added sugars) and the risk of type 2 diabetes mellitus.

4.3. Conclusions and recommendations

This umbrella review shows that high dietary sugar consumption, especially intake of sugars that contain fructose, is harmfully associated with large numbers of health outcomes.

Evidence for the harmful associations between dietary sugar consumption and changes in body weight (sugar sweetened beverages), ectopic fat accumulation (added sugars), obesity in children (sugar sweetened beverages), coronary heart disease (sugar sweetened beverages), and depression (sugar sweetened beverages) seems to be more reliable than that for other outcomes. Evidence of the association between dietary sugar consumption and cancer remains limited but warrants further research. In combination with the WHO and WCRF/AICR recommendations and our findings, we recommend reducing the consumption of free sugars or added sugars to below 25 g/day (approximately six teaspoons a day) and limiting the consumption of sugar sweetened beverages to less than one serving a week (approximately 200-355 mL/week).[38 119] To change sugar consumption patterns, especially for children and adolescents, a combination of widespread public health education and policies worldwide is urgently needed.

5. What is already known on this topic

Sugar consumption could have negative effects on health, especially obesity, diabetes, cardiovascular disease, hyperuricaemia, gout, ectopic fatty accumulation, dental caries, and some cancers

Deficiencies in study design, varying measurements, inconsistent findings, and different definitions of exposure make drawing final conclusions on associations difficult Comprehensive evaluation of the quality of existing evidence on the associations of sugar consumption with all health outcomes is needed

6. What this study adds

High dietary sugar consumption is generally more harmful than beneficial for health, especially in cardiometabolic disease

Evidence of the association between dietary sugar consumption and cancer remains limited but warrants further research

Existing evidence is mostly observational and of low quality, and further randomised controlled trials are needed

7. Acknowledgments

We thank Nanxi Yan for her linguistic assistance during the preparation and revision of this manuscript.

8. Web extra

Extra material supplied by authors

Web appendix: Supplementary materials

9. Notes

Contributors: YH, ZYC, BC, and, JZL are joint first authors and contributed equally to this work. QW, DHC, and LRL are joint corresponding authors and contributed equally to this work. YH, ZYC, BC, and JZL conducted study selection, data extraction, and analysis and wrote the manuscript. QW, DHC, and LRL designed the study, supervised the project, and revised the manuscript. XY, JL, WW, TTD, HYC, YW, RYW, PZW, JBG, QD, and CFL assisted with detailed statistical analysis. All authors reviewed and approved the final version of the manuscript. LRL is the guarantor. The corresponding authors

(QW, DHC, and LRL) attest that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study was funded by the National Natural Science Foundation of China (grant number 82000721) and Program from the Department of Science and Technology of Sichuan Province (grant number 2020YJ0054). The funders had no role in considering the study design or in the collection, analysis, and interpretation of data, the writing of the report, or the decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from the National Natural Science Foundation of China and Program from the Department of Science and Technology of Sichuan Province for the submitted work; no financial relationship with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been ex- plained.

Dissemination to participants and related patients and public communities: After publication, the findings of this review will be disseminated to appropriate audiences such as academia, clinicians, policy makers, and the general public, through various channels including blogs, press releases, and social media.

Provenance and peer review: Not commissioned; externally peer reviewed.

10. Ethics statements

10.1. Ethical approval

Not needed.

11. Data availability statement

The list of all meta-analyses not selected for data extraction and reanalysis is available if needed.

12. References

1. Qin P, Li Q, Zhao Y, *et al.*. Sugar and artificially sweetened beverages and risk of obesity, type 2 diabetes mellitus, hypertension, and all-cause mortality: a dose-response meta-analysis of prospective cohort studies. *Eur J Epidemiol* 2020;35:655-71. 10.1007/s10654-020-00655-y [PubMed: 32529512] [CrossRef: 10.1007/s10654-020-00655-y]
2. Ruanpeng D, Thongprayoon C, Cheungpasitporn W, Harindhanavudhi T. Sugar and artificially sweetened beverages linked to obesity: a systematic review and meta-analysis. *QJM* 2017;110:513-20. 10.1093/qjmed/hcx068 [PubMed: 28402535] [CrossRef: 10.1093/qjmed/hcx068]
3. Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. *BMJ* 2012;346:e7492. 10.1136/bmj.e7492 [PubMed: 23321486] [CrossRef: 10.1136/bmj.e7492]
4. Malik VS, Popkin BM, Bray GA, Després JP, Willett WC, Hu FB. Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: a meta-analysis. *Diabetes Care* 2010;33:2477-83. 10.2337/dc10-1079 [PMCID: PMC2963518] [PubMed: 20693348] [CrossRef: 10.2337/dc10-1079]
5. Imamura F, O'Connor L, Ye Z, *et al.*. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. *BMJ* 2015;351:h3576. 10.1136/bmj.h3576 [PMCID: PMC4510779] [PubMed: 26199070] [CrossRef: 10.1136/bmj.h3576]
6. Meng Y, Li S, Khan J, *et al.*. Sugar- and Artificially Sweetened Beverages Consumption Linked to Type 2 Diabetes, Cardiovascular Diseases, and All-Cause Mortality: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. *Nutrients* 2021;13:2636. 10.3390/nu13082636 [PMCID: PMC8402166] [PubMed: 34444794] [CrossRef: 10.3390/nu13082636]
7. Xi B, Huang Y, Reilly KH, *et al.*. Sugar-sweetened beverages and risk of hypertension and CVD: a dose-response meta-analysis. *Br J Nutr* 2015;113:709-17. 10.1017/S0007114514004383 [PubMed: 25735740] [CrossRef: 10.1017/S0007114514004383]
8. Yin J, Zhu Y, Malik V, *et al.*. Intake of Sugar-Sweetened and Low-Calorie Sweetened Beverages and Risk of Cardiovascular Disease: A Meta-Analysis and Systematic Review. *Adv Nutr* 2021;12:89-101. 10.1093/advances/nmaa084 [PMCID: PMC7850046] [PubMed: 32696948] [CrossRef: 10.1093/advances/nmaa084]
9. Huang C, Huang J, Tian Y, Yang X, Gu D. Sugar sweetened beverages consumption and risk of coronary heart disease: a meta-analysis of prospective studies. *Atherosclerosis* 2014;234:11-6. 10.1016/j.atherosclerosis.2014.01.037 [PubMed: 24583500] [CrossRef: 10.1016/j.atherosclerosis.2014.01.037]

10. Bechthold A, Boeing H, Schwedhelm C, *et al.*. Food groups and risk of coronary heart disease, stroke and heart failure: A systematic review and dose-response meta-analysis of prospective studies. *Crit Rev Food Sci Nutr* 2019;59:1071-90. 10.1080/10408398.2017.1392288 [PubMed: 29039970] [CrossRef: 10.1080/10408398.2017.1392288]
11. Ebrahimpour-Koujan S, Saneei P, Larijani B, Esmaillzadeh A. Consumption of sugar sweetened beverages and dietary fructose in relation to risk of gout and hyperuricemia: a systematic review and meta-analysis. *Crit Rev Food Sci Nutr* 2020;60:1-10. 10.1080/10408398.2018.1503155 [PubMed: 30277800] [CrossRef: 10.1080/10408398.2018.1503155]
12. Lee Y, Song GG. Association between Sugar-Sweetened Beverage Consumption and the Risk of Gout: A Meta- Analysis. *J Rheum Dis* 2016;23:304-10 10.4078/jrd.2016.23.5.304. [CrossRef: 10.4078/jrd.2016.23.5.304]
13. Ayoub-Charette S, Liu Q, Khan TA, *et al.*. Important food sources of fructose-containing sugars and incident gout: a systematic review and meta-analysis of prospective cohort studies. *BMJ Open* 2019;9:e024171. 10.1136/bmjopen- 2018-024171 [PMCID: PMC6502023] [PubMed: 31061018] [CrossRef: 10.1136/bmjopen-2018-024171]
14. Wijarnpreecha K, Thongprayoon C, Edmonds PJ, Cheungpasitporn W. Associations of sugar- and artificially sweetened soda with nonalcoholic fatty liver disease: a systematic review and meta-analysis. *QJM* 2016;109:461-6. 10.1093/qjmed/hcv172 [PubMed: 26385233] [CrossRef: 10.1093/qjmed/hcv172]
15. Asgari-Taee F, Zerafati-Shoae N, Dehghani M, Sadeghi M, Baradaran HR, Jazayeri S. Association of sugar sweetened beverages consumption with non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Eur J Nutr* 2019;58:1759-69. 10.1007/s00394-018-1711-4 [PubMed: 29761318] [CrossRef: 10.1007/s00394-018-1711-4]
16. Chen H, Wang J, Li Z, *et al.*. Consumption of sugar-sweetened beverages has a dose-dependent effect on the risk of non-alcoholic fatty liver disease: An updated systematic review and dose-response meta-analysis. *Int J Environ Res Public Health* 2019;16:2192. 10.3390/ijerph16122192 [PMCID: PMC6617076] [PubMed: 31234281] [CrossRef: 10.3390/ijerph16122192]
17. Valenzuela MJ, Waterhouse B, Aggarwal VR, Bloor K, Doran T. Effect of sugar-sweetened beverages on oral health: a systematic review and meta-analysis. *Eur J Public Health* 2021;31:122-9. 10.1093/eurpub/ckaa147 [PubMed: 32830237] [CrossRef: 10.1093/eurpub/ckaa147]
18. Li Y, Guo L, He K, Huang C, Tang S. Consumption of sugar-sweetened beverages and fruit juice and human cancer: a systematic review and dose-response meta-analysis of observational studies. *J Cancer* 2021;12:3077-88. 10.7150/jca.51322 [PMCID: PMC8040874] [PubMed: 33854607] [CrossRef: 10.7150/jca.51322]
19. Llaha F, Gil-Lespinard M, Unal P, de Villasante I, Castañ eda J, Zamora-Ros R. Consumption of Sweet Beverages and Cancer Risk. A Systematic Review and Meta-Analysis of Observational Studies. *Nutrients* 2021;13:516. 10.3390/nu13020516 [PMCID: PMC7915548] [PubMed: 33557387] [CrossRef: 10.3390/nu13020516]
20. Hur J, Otegbeye E, Joh HK, *et al.*. Sugar-sweetened beverage intake in adulthood and adolescence and risk of early- onset colorectal cancer among women. *Gut* 2021;70:2330-6. 10.1136/gutjnl-2020-323450 [PMCID: PMC8571123] [PubMed: 33958435] [CrossRef: 10.1136/gutjnl-2020-323450]
21. Chazelas E, Srour B, Desmetz E, *et al.*. Sugary drink consumption and risk of cancer: results from NutriNet-Santé prospective cohort. *BMJ* 2019;366:l2408. 10.1136/bmj.l2408 [PMCID: PMC6614796] [PubMed: 31292122] [CrossRef: 10.1136/bmj.l2408]
22. Mann J, Cummings JH, Englyst HN, *et al.*. FAO/WHO scientific update on carbohydrates in human nutrition: conclusions. *Eur J Clin Nutr* 2007;61(Suppl 1):S132-7. 10.1038/sj.ejcn.1602943 [PubMed: 17992184] [CrossRef: 10.1038/sj.ejcn.1602943]
23. Phillips JA. Dietary Guidelines for Americans, 2020-2025. *Workplace Health Saf* 2021;69:395. 10.1177/21650799211026980 [PubMed: 34279148] [CrossRef: 10.1177/21650799211026980]
24. Malik VS, Schulze MB, Hu FB. Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr* 2006;84:274-88. 10.1093/ajcn/84.2.274 [PMCID: PMC3210834] [PubMed: 16895873] [CrossRef: 10.1093/ajcn/84.2.274]
25. Rosinger A, Herrick K, Gahche J, Park S. Sugar-sweetened Beverage Consumption Among U.S. Youth, 2011-2014. *NCHS Data Brief* 2017;(271):1-8. [PubMed: 28135184]
26. Blecher E, Liber AC, Drope JM, Nguyen B, Stoklosa M. Global Trends in the Affordability of Sugar-Sweetened Beverages, 1990-2016. *Prev Chronic Dis* 2017;14:E37. 10.5888/pcd14.160406 [PMCID: PMC5420443] [PubMed: 28472607] [CrossRef: 10.5888/pcd14.160406]
27. Malik VS, Hu FB. The role of sugar-sweetened beverages in the global epidemics of obesity and chronic diseases. *Nat Rev Endocrinol* 2022;18:205-18. 10.1038/s41574-021-00627-6 [PMCID: PMC8778490] [PubMed: 35064240] [CrossRef: 10.1038/s41574-021-00627-6]
28. Welsh JA, Sharma AJ, Grellinger L, Vos MB. Consumption of added sugars is decreasing in the United States. *Am J Clin Nutr* 2011;94:726-34. 10.3945/ajcn.111.018366 [PMCID: PMC3155936] [PubMed: 21753067] [CrossRef: 10.3945/ajcn.111.018366]
29. Malik VS, Willett WC, Hu FB. Global obesity: trends, risk factors and policy implications. *Nat Rev Endocrinol* 2013;9:13-27. 10.1038/nrendo.2012.199 [PubMed: 23165161] [CrossRef: 10.1038/nrendo.2012.199]
30. El-Malky AM, Naik RG, Elnouman AA. Sugary beverages consumption and latent autoimmune diabetes in adults: Systematic review and meta-analysis. *Clin Diabetol* 2020;9:118-27 10.5603/DK.2020.0007. [CrossRef: 10.5603/DK.2020.0007]
31. Qin Z, Xu F, Ye Q, *et al.*. Sugar-sweetened beverages and school students' hypertension in urban areas of Nanjing, China. *J Hum Hypertens* 2018;32:392-6. 10.1038/s41371-018-0030-9 [PubMed: 29410454] [CrossRef: 10.1038/s41371-018-0030-9]
32. Kit BK, Fakhouri TH, Park S, Nielsen SJ, Ogden CL. Trends in sugar-sweetened beverage consumption among youth and adults in the United States: 1999-2010. *Am J Clin Nutr* 2013;98:180-8. 10.3945/ajcn.112.057943 [PubMed: 23676424] [CrossRef: 10.3945/ajcn.112.057943]

33. Singh GM, Micha R, Khatibzadeh S, *et al.*Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (NutriCoDE) . Global, Regional, and National Consumption of Sugar-Sweetened Beverages, Fruit Juices, and Milk: A Systematic Assessment of Beverage Intake in 187 Countries. *PLoS One* 2015;10:e0124845. 10.1371/journal.pone.0124845 [PMCID: PMC4526649] [PubMed: 26244332] [CrossRef: 10.1371/journal.pone.0124845]
34. Colchero MA, Popkin BM, Rivera JA, Ng SW. Beverage purchases from stores in Mexico under the excise tax on sugar sweetened beverages: observational study. *BMJ* 2016;352:h6704. 10.1136/bmj.h6704 [PMCID: PMC4986313] [PubMed: 26738745] [CrossRef: 10.1136/bmj.h6704]
35. Taillie LS, Reyes M, Colchero MA, Popkin B, Corvalán C. An evaluation of Chile's Law of Food Labeling and Advertising on sugar-sweetened beverage purchases from 2015 to 2017: A before-and-after study. *PLoS Med* 2020;17:e1003015. 10.1371/journal.pmed.1003015 [PMCID: PMC7012389] [PubMed: 32045424] [CrossRef: 10.1371/journal.pmed.1003015]
36. Basu S, Vellakkal S, Agrawal S, Stuckler D, Popkin B, Ebrahim S. Averting obesity and type 2 diabetes in India through sugar-sweetened beverage taxation: an economic-epidemiologic modeling study. *PLoS Med* 2014;11:e1001582. 10.1371/journal.pmed.1001582 [PMCID: PMC3883641] [PubMed: 24409102] [CrossRef: 10.1371/journal.pmed.1001582]
37. Manyema M, Veerman LJ, Chola L, *et al.*. The potential impact of a 20% tax on sugar-sweetened beverages on obesity in South African adults: a mathematical model. *PLoS One* 2014;9:e105287. 10.1371/journal.pone.0105287 [PMCID: PMC4138175] [PubMed: 25136987] [CrossRef: 10.1371/journal.pone.0105287]
38. World Health Organization . *Guideline: Sugars Intake for Adults and Children*. World Health Organization, 2015. [PubMed: 25905159]
39. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Based Healthc* 2015;13:132-40. 10.1097/XEB.0000000000000055 [PubMed: 26360830] [CrossRef: 10.1097/XEB.0000000000000055]
40. Papatheodorou S. Umbrella reviews: what they are and why we need them. *Eur J Epidemiol* 2019;34:543-6. 10.1007/s10654-019-00505-6 [PubMed: 30852716] [CrossRef: 10.1007/s10654-019-00505-6]
41. Scottish Intercollegiate Guidelines Network. Search Filters. https://www.sign.ac.uk/what-we-do/methodology/search-filters/.
42. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *BMJ* 2017;359:j5024. 10.1136/bmj.j5024 [PMCID: PMC5696634] [PubMed: 29167102] [CrossRef: 10.1136/bmj.j5024]
43. Huang Y, Cao D, Chen Z, *et al.*. Iron intake and multiple health outcomes: Umbrella review. *Crit Rev Food Sci Nutr* 2021;1-18. 10.1080/10408398.2021.1982861 [PubMed: 34583608] [CrossRef: 10.1080/10408398.2021.1982861]
44. Shea BJ, Grimshaw JM, Wells GA, *et al.*. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10. 10.1186/1471-2288-7-10 [PMCID: PMC1810543] [PubMed: 17302989] [CrossRef: 10.1186/1471-2288-7-10]
45. Guyatt G, Oxman AD, Akl EA, *et al.*. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383-94. 10.1016/j.jclinepi.2010.04.026 [PubMed: 21195583] [CrossRef: 10.1016/j.jclinepi.2010.04.026]
46. Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 2009;181:488-93. 10.1503/cmaj.081086 [PMCID: PMC2761440] [PubMed: 19654195] [CrossRef: 10.1503/cmaj.081086]
47. Veronese N, Solmi M, Caruso MG, *et al.*. Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. *Am J Clin Nutr* 2018;107:436-44. 10.1093/ajcn/nqx082 [PubMed: 29566200] [CrossRef: 10.1093/ajcn/nqx082]
48. Wallace TC, Bailey RL, Blumberg JB, *et al.*. Fruits, vegetables, and health: A comprehensive narrative, umbrella review of the science and recommendations for enhanced public policy to improve intake. *Crit Rev Food Sci Nutr* 2020;60:2174-211. 10.1080/10408398.2019.1632258 [PubMed: 31267783] [CrossRef: 10.1080/10408398.2019.1632258]
49. Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. *BMJ* 2014;348:g2035. 10.1136/bmj.g2035 [PMCID: PMC3972415] [PubMed: 24690624] [CrossRef: 10.1136/bmj.g2035]
50. Huang Y, Cao D, Chen Z, *et al.*. Red and processed meat consumption and cancer outcomes: Umbrella review. *Food Chem* 2021;356:129697. 10.1016/j.foodchem.2021.129697 [PubMed: 33838606] [CrossRef: 10.1016/j.foodchem.2021.129697]
51. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34. 10.1136/bmj.315.7109.629 [PMCID: PMC2127453] [PubMed: 9310563] [CrossRef: 10.1136/bmj.315.7109.629]
52. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74. 10.2307/2529310 [PubMed: 843571] [CrossRef: 10.2307/2529310]
53. Malik VS, Pan A, Willett WC, Hu FB. Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *Am J Clin Nutr* 2013;98:1084-102. 10.3945/ajcn.113.058362 [PMCID: PMC3778861] [PubMed: 23966427] [CrossRef: 10.3945/ajcn.113.058362]
54. Ma J, Karlsen MC, Chung M, *et al.*. Potential link between excess added sugar intake and ectopic fat: a systematic review of randomized controlled trials. *Nutr Rev* 2016;74:18-32. 10.1093/nutrit/nuv047 [PMCID: PMC4859325] [PubMed: 26518034] [CrossRef: 10.1093/nutrit/nuv047]
55. Wang DD, Sievenpiper JL, de Souza RJ, *et al.*. The effects of fructose intake on serum uric acid vary among controlled die-

tary trials. *J Nutr* 2012;142:916-23. 10.3945/jn.111.151951 [PMCID: PMC3327749] [PubMed: 22457397] [CrossRef: 10.3945/jn.111.151951]

56. Sievenpiper JL, de Souza RJ, Mirrahimi A, *et al.*. Effect of fructose on body weight in controlled feeding trials: a systematic review and meta-analysis. *Ann Intern Med* 2012;156:291-304.10.7326/0003-4819-156-4-201202210-00007 [PubMed: 22351714] [CrossRef: 10.7326/0003-4819-156-4-201202210-00007]
57. Ebrahimpour-Koujan S, Saneei P, Larijani B, Esmaillzadeh A. Consumption of sugar-sweetened beverages and serum uric acid concentrations: a systematic review and meta-analysis. *J Hum Nutr Diet* 2021;34:305-13. 10.1111/jhn.12796 [PubMed: 32683776] [CrossRef: 10.1111/jhn.12796]
58. Jamnik J, Rehman S, Blanco Mejia S, *et al.*. Fructose intake and risk of gout and hyperuricemia: a systematic review and meta-analysis of prospective cohort studies. *BMJ Open* 2016;6:e013191. 10.1136/bmjopen-2016-013191 [PMCID: PMC5073537] [PubMed: 27697882] [CrossRef: 10.1136/bmjopen-2016-013191]
59. Li R, Yu K, Li C. Dietary factors and risk of gout and hyperuricemia: a meta-analysis and systematic review. *Asia Pac J Clin Nutr* 2018;27:1344-56. [PubMed: 30485934]
60. Nikniaz L, Abbasalizad-Farhangi M, Vajdi M, Nikniaz Z. The association between Sugars Sweetened Beverages (SSBs) and lipid profile among children and youth: A systematic review and dose-response meta-analysis of cross- sectional studies. *Pediatr Obes* 2021;16:e12782. 10.1111/ijpo.12782 [PubMed: 33629539] [CrossRef: 10.1111/ijpo.12782]
61. Semnani-Azad Z, Khan TA, Blanco Mejia S, *et al.*. Association of Major Food Sources of Fructose-Containing Sugars With Incident Metabolic Syndrome: A Systematic Review and Meta-analysis. *JAMA Netw Open* 2020;3:e209993. 10.1001/jamanetworkopen.2020.9993 [PMCID: PMC7348689] [PubMed: 32644139] [CrossRef: 10.1001/jamanetworkopen.2020.9993]
62. Tsilas CS, de Souza RJ, Mejia SB, *et al.*. Relation of total sugars, fructose and sucrose with incident type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies. *CMAJ* 2017;189:E711-20. 10.1503/cmaj.160706 [PMCID: PMC5436961] [PubMed: 28536126] [CrossRef: 10.1503/cmaj.160706]
63. Mattes RD, Shikany JM, Kaiser KA, Allison DB. Nutritively sweetened beverage consumption and body weight: a systematic review and meta-analysis of randomized experiments. *Obes Rev* 2011;12:346-65.10.1111/j.1467-789X.2010.00755.x [PMCID: PMC3169649] [PubMed: 20524996] [CrossRef: 10.1111/j.1467-789X.2010.00755.x]
64. Ardeshirlarijani E, Jalilpiran Y, Daneshzad E, Larijani B, Namazi N, Azadbakht L. Association between sugar- sweetened beverages and waist circumference in adult populations: A meta-analysis of prospective cohort studies. *Clin Nutr ESPEN* 2021;41:118-25. 10.1016/j.clnesp.2020.10.014 [PubMed: 33487253] [CrossRef: 10.1016/j.clnesp.2020.10.014]
65. David Wang D, Sievenpiper JL, de Souza RJ, *et al.*. Effect of fructose on postprandial triglycerides: a systematic review and meta-analysis of controlled feeding trials. *Atherosclerosis* 2014;232:125-33. 10.1016/j.atherosclerosis.2013.10.019 [PubMed: 24401226] [CrossRef: 10.1016/j.atherosclerosis.2013.10.019]
66. Narain A, Kwok CS, Mamas MA. Soft drinks and sweetened beverages and the risk of cardiovascular disease and mortality: a systematic review and meta-analysis. *Int J Clin Pract* 2016;70:791-805. 10.1111/ijcp.12841 [PubMed: 27456347] [CrossRef: 10.1111/ijcp.12841]
67. Jayalath VH, Sievenpiper JL, de Souza RJ, *et al.*. Total fructose intake and risk of hypertension: a systematic review and meta-analysis of prospective cohorts. *J Am Coll Nutr* 2014;33:328-39. 10.1080/07315724.2014.916237 [PMCID: PMC4261182] [PubMed: 25144126] [CrossRef: 10.1080/07315724.2014.916237]
68. Zhang YB, Jiang YW, Chen JX, Xia PF, Pan A. Association of Consumption of Sugar-Sweetened Beverages or Artificially Sweetened Beverages with Mortality: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. *Adv Nutr* 2021;12:374-83. 10.1093/advances/nmaa110 [PMCID: PMC8009739] [PubMed: 33786594] [CrossRef: 10.1093/advances/nmaa110]
69. Farhangi MA, Nikniaz L, Khodarahmi M. Sugar-sweetened beverages increases the risk of hypertension among children and adolescence: a systematic review and dose-response meta-analysis. *J Transl Med* 2020;18:344. 10.1186/s12967-020-02511-9 [PMCID: PMC7487688] [PubMed: 32891165] [CrossRef: 10.1186/s12967-020-02511-9]
70. Liu Q, Ayoub-Charette S, Khan TA, *et al.*. Important Food Sources of Fructose-Containing Sugars and Incident Hypertension: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. *J Am Heart Assoc* 2019;8:e010977. 10.1161/JAHA.118.010977 [PMCID: PMC6951071] [PubMed: 31826724] [CrossRef: 10.1161/JAHA.118.010977]
71. Khan TA, Tayyiba M, Agarwal A, *et al.*. Relation of Total Sugars, Sucrose, Fructose, and Added Sugars With the Risk of Cardiovascular Disease: A Systematic Review and Dose-Response Meta-analysis of Prospective Cohort Studies. *Mayo Clin Proc* 2019;94:2399-414. 10.1016/j.mayocp.2019.05.034 [PubMed: 31806098] [CrossRef: 10.1016/j.mayocp.2019.05.034]
72. Aune D, Chan DSM, Vieira AR, *et al.*. Dietary fructose, carbohydrates, glycemic indices and pancreatic cancer risk: a systematic review and meta-analysis of cohort studies. *Ann Oncol* 2012;23:2536-46. 10.1093/annonc/mds076 [PubMed: 22539563] [CrossRef: 10.1093/annonc/mds076]
73. Zhang X, Albanes D, Beeson WL, *et al.*. Risk of colon cancer and coffee, tea, and sugar-sweetened soft drink intake: pooled analysis of prospective cohort studies. *J Natl Cancer Inst* 2010;102:771-83. 10.1093/jnci/djq107 [PMCID: PMC2879415] [PubMed: 20453203] [CrossRef: 10.1093/jnci/djq107]
74. Galeone C, Pelucchi C, La Vecchia C. Added sugar, glycemic index and load in colon cancer risk. *Curr Opin Clin Nutr Metab Care* 2012;15:368-73. 10.1097/MCO.0b013e3283539f81 [PubMed: 22510682] [CrossRef: 10.1097/MCO.0b013e3283539f81]
75. Liu J, Tang W, Sang L, *et al.*. Milk, yogurt, and lactose intake and ovarian cancer risk: a meta-analysis. *Nutr Cancer* 2015;67:68-72. 10.1080/01635581.2014.956247 [PubMed: 25298278] [CrossRef: 10.1080/01635581.2014.956247]

76. Pan B, Ge L, Lai H, *et al.*. Association of soft drink and 100% fruit juice consumption with all-cause mortality, cardiovascular diseases mortality, and cancer mortality: A systematic review and dose-response meta-analysis of prospective cohort studies. *Crit Rev Food Sci Nutr* 2022;62:8908-19. 10.1080/10408398.2021.1937040 [PubMed: 34121531] [CrossRef: 10.1080/10408398.2021.1937040]

77. Hu D, Cheng L, Jiang W. Sugar-sweetened beverages consumption and the risk of depression: A meta-analysis of observational studies. *J Affect Disord* 2019;245:348-55. 10.1016/j.jad.2018.11.015 [PubMed: 30419536] [CrossRef: 10.1016/j.jad.2018.11.015]

78. Chiu S, Sievenpiper JL, de Souza RJ, *et al.*. Effect of fructose on markers of non-alcoholic fatty liver disease (NAFLD): a systematic review and meta-analysis of controlled feeding trials. *Eur J Clin Nutr* 2014;68:416-23. 10.1038/ejcn.2014.8 [PMCID: PMC3975811] [PubMed: 24569542] [CrossRef: 10.1038/ejcn.2014.8]

79. Al-Zalabani AH, Noor Elahi I, Katib A, *et al.*. Association between soft drinks consumption and asthma: a systematic review and meta-analysis. *BMJ Open* 2019;9:e029046. 10.1136/bmjopen-2019-029046 [PMCID: PMC6797295] [PubMed: 31615794] [CrossRef: 10.1136/bmjopen-2019-029046]

80. Farsad-Naeimi A, Asjodi F, Omidian M, *et al.*. Sugar consumption, sugar sweetened beverages and Attention Deficit Hyperactivity Disorder: A systematic review and meta-analysis. *Complement Ther Med* 2020;53:102512. 10.1016/j.ctim.2020.102512 [PubMed: 33066852] [CrossRef: 10.1016/j.ctim.2020.102512]

81. Ahn H, Park YK. Sugar-sweetened beverage consumption and bone health: a systematic review and meta- analysis. *Nutr J* 2021;20:41. 10.1186/s12937-021-00698-1 [PMCID: PMC8101184] [PubMed: 33952276] [CrossRef: 10.1186/s12937-021-00698-1]

82. Lo WC, Ou SH, Chou CL, Chen JS, Wu MY, Wu MS. Sugar- and artificially-sweetened beverages and the risks of chronic kidney disease: a systematic review and dose-response meta-analysis. *J Nephrol* 2021;34:1791-804. 10.1007/s40620-020-00957-0 [PubMed: 33502726] [CrossRef: 10.1007/s40620-020-00957-0]

83. Gupta A, Singh A, Fernando RL, Dharmage SC, Lodge CJ, Waidyatillake NT. The association between sugar intake during pregnancy and allergies in offspring: a systematic review and a meta-analysis of cohort studies. *Nutr Rev* 2022;80:904-18. 10.1093/nutrit/nuab052 [PubMed: 34432049] [CrossRef: 10.1093/nutrit/nuab052]

84. Brisbois TD, Marsden SL, Anderson GH, Sievenpiper JL. Estimated intakes and sources of total and added sugars in the Canadian diet. *Nutrients* 2014;6:1899-912. 10.3390/nu6051899 [PMCID: PMC4042566] [PubMed: 24815507] [CrossRef: 10.3390/nu6051899]

85. Marriott BP, Cole N, Lee E. National estimates of dietary fructose intake increased from 1977 to 2004 in the United States. *J Nutr* 2009;139:1228S-35S. 10.3945/jn.108.098277 [PubMed: 19403716] [CrossRef: 10.3945/jn.108.098277]

86. Bernstein JT, Lou W, L'Abbe MR. Examining the Relationship between Free Sugars and Calorie Contents in Canadian Prepacked Foods and Beverages. *Foods* 2017;6:75. 10.3390/foods6090075 [PMCID: PMC5615287] [PubMed: 28872586] [CrossRef: 10.3390/foods6090075]

87. Aune D, Norat T, Romundstad P, Vatten LJ. Dairy products and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. *Am J Clin Nutr* 2013;98:1066-83. 10.3945/ajcn.113.059030 [PubMed: 23945722] [CrossRef: 10.3945/ajcn.113.059030]

88. Aune D, Norat T, Romundstad P, Vatten LJ. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. *Eur J Epidemiol* 2013;28:845-58. 10.1007/s10654-013-9852-5 [PubMed: 24158434] [CrossRef: 10.1007/s10654-013-9852-5]

89. Li M, Fan Y, Zhang X, Hou W, Tang Z. Fruit and vegetable intake and risk of type 2 diabetes mellitus: meta-analysis of prospective cohort studies. *BMJ Open* 2014;4:e005497. 10.1136/bmjopen-2014-005497 [PMCID: PMC4225228] [PubMed: 25377009] [CrossRef: 10.1136/bmjopen-2014-005497]

90. Zhan J, Liu YJ, Cai LB, Xu FR, Xie T, He QQ. Fruit and vegetable consumption and risk of cardiovascular disease: A meta-analysis of prospective cohort studies. *Crit Rev Food Sci Nutr* 2017;57:1650-63. 10.1080/10408398.2015.1008980 [PubMed: 26114864] [CrossRef: 10.1080/10408398.2015.1008980]

91. Aune D, Keum N, Giovannucci E, *et al.*. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: systematic review and dose-response meta-analysis of prospective studies. *BMJ* 2016;353:i2716. 10.1136/bmj.i2716 [PMCID: PMC4908315] [PubMed: 27301975] [CrossRef: 10.1136/bmj.i2716]

92. Alexander DD, Bylsma LC, Vargas AJ, *et al.*. Dairy consumption and CVD: a systematic review and meta-analysis. *Br J Nutr* 2016;115:737-50. 10.1017/S0007114515005000 [PubMed: 26786887] [CrossRef: 10.1017/S0007114515005000]

93. Janevski M, Ratnayake S, Siljanovski S, McGlynn MA, Cameron-Smith D, Lewandowski P. Fructose containing sugars modulate mRNA of lipogenic genes ACC and FAS and protein levels of transcription factors ChREBP and SREBP1c with no effect on body weight or liver fat. *Food Funct* 2012;3:141-9. 10.1039/C1FO10111K [PubMed: 22159273] [CrossRef: 10.1039/C1FO10111K]

94. Kim MS, Krawczyk SA, Doridot L, *et al.*. ChREBP regulates fructose-induced glucose production independently of insulin signaling. *J Clin Invest* 2016;126:4372-86. 10.1172/JCI81993 [PMCID: PMC5096918] [PubMed: 27669460] [CrossRef: 10.1172/JCI81993]

95. Koo HY, Wallig MA, Chung BH, Nara TY, Cho BH, Nakamura MT. Dietary fructose induces a wide range of genes with distinct shift in carbohydrate and lipid metabolism in fed and fasted rat liver. *Biochim Biophys Acta* 2008;1782:341-8. 10.1016/j.bbadis.2008.02.007 [PubMed: 18346472] [CrossRef: 10.1016/j.bbadis.2008.02.007]

96. Zhao S, Jang C, Liu J, *et al.*. Dietary fructose feeds hepatic lipogenesis via microbiota-derived acetate. *Nature* 2020;579:586-91. 10.1038/s41586-020-2101-7 [PMCID: PMC7416516] [PubMed: 32214246] [CrossRef: 10.1038/s41586-020-2101-7]

97. Badin PM, Louche K, Mairal A, *et al.*. Altered skeletal muscle lipase expression and activity contribute to insulin resistance in humans. *Diabetes* 2011;60:1734-42. 10.2337/db10-1364 [PMCID: PMC3114384] [PubMed: 21498783] [CrossRef: 10.2337/db10-1364]

98. Kumashiro N, Erion DM, Zhang D, *et al.*. Cellular mechanism of insulin resistance in nonalcoholic fatty liver disease. *Proc Natl Acad Sci U S A* 2011;108:16381-5. 10.1073/pnas.1113359108 [PMCID: PMC3182681] [PubMed: 21930939] [CrossRef: 10.1073/pnas.1113359108]

99. Petersen MC, Shulman GI. Roles of Diacylglycerols and Ceramides in Hepatic Insulin Resistance. *Trends Pharmacol Sci* 2017;38:649-65. 10.1016/j.tips.2017.04.004 [PMCID: PMC5499157] [PubMed: 28551355] [CrossRef: 10.1016/j.tips.2017.04.004]

100. Softic S, Meyer JG, Wang GX, *et al.*. Dietary Sugars Alter Hepatic Fatty Acid Oxidation via Transcriptional and Post-translational Modifications of Mitochondrial Proteins. *Cell Metab* 2019;30:735-753.e4. 10.1016/j.cmet.2019.09.003 [PMCID: PMC7816129] [PubMed: 31577934] [CrossRef: 10.1016/j.cmet.2019.09.003]

101. Taylor SR, Ramsamooj S, Liang RJ, *et al.*. Dietary fructose improves intestinal cell survival and nutrient absorption. *Nature* 2021;597:263-7. 10.1038/s41586-021-03827-2 [PMCID: PMC8686685] [PubMed: 34408323] [CrossRef: 10.1038/s41586-021-03827-2]

102. Cox CL, Stanhope KL, Schwarz JM, *et al.*. Consumption of fructose-sweetened beverages for 10 weeks reduces net fat oxidation and energy expenditure in overweight/obese men and women. *Eur J Clin Nutr* 2012;66:201-8. 10.1038/ejcn.2011.159 [PMCID: PMC3252467] [PubMed: 21952692] [CrossRef: 10.1038/ejcn.2011.159]

103. Sundborn G, Thornley S, Merriman TR, *et al.*. Are Liquid Sugars Different from Solid Sugar in Their Ability to Cause Metabolic Syndrome? *Obesity (Silver Spring)* 2019;27:879-87. 10.1002/oby.22472 [PubMed: 31054268] [CrossRef: 10.1002/oby.22472]

104. Malik VS. Sugar sweetened beverages and cardiometabolic health. *Curr Opin Cardiol* 2017;32:572-9. 10.1097/HCO.0000000000000439 [PubMed: 28639973] [CrossRef: 10.1097/HCO.0000000000000439]

105. DellaValle DM, Roe LS, Rolls BJ. Does the consumption of caloric and non-caloric beverages with a meal affect energy intake? *Appetite* 2005;44:187-93. 10.1016/j.appet.2004.11.003 [PubMed: 15808893] [CrossRef: 10.1016/j.appet.2004.11.003]

106. Raben A, Vasilaras TH, Møller AC, Astrup A. Sucrose compared with artificial sweeteners: different effects on ad libitum food intake and body weight after 10 wk of supplementation in overweight subjects. *Am J Clin Nutr* 2002;76:721-9. 10.1093/ajcn/76.4.721 [PubMed: 12324283] [CrossRef: 10.1093/ajcn/76.4.721]

107. Reid M, Hammersley R, Hill AJ, Skidmore P. Long-term dietary compensation for added sugar: effects of supplementary sucrose drinks over a 4-week period. *Br J Nutr* 2007;97:193-203. 10.1017/S0007114507252705 [PubMed: 17217576] [CrossRef: 10.1017/S0007114507252705]

108. Geidl-Flueck B, Hochuli M, Németh A, , *et al.*. Fructose- and sucrose- but not glucose-sweetened beverages promote hepatic de novo lipogenesis: A randomized controlled trial. *J Hepatol* 2021;75:46-54. 10.1016/j.jhep.2021.02.027 [PubMed: 33684506] [CrossRef: 10.1016/j.jhep.2021.02.027]

109. Choi HK, Curhan G. Soft drinks, fructose consumption, and the risk of gout in men: prospective cohort study. *BMJ* 2008;336:309-12. 10.1136/bmj.39449.819271.BE [PMCID: PMC2234536] [PubMed: 18244959] [CrossRef: 10.1136/bmj.39449.819271.BE]

110. Choi HK, Willett W, Curhan G. Fructose-rich beverages and risk of gout in women. *JAMA* 2010;304:2270-8. 10.1001/jama.2010.1638 [PMCID: PMC3058904] [PubMed: 21068145] [CrossRef: 10.1001/jama.2010.1638]

111. Kanbay M, Jensen T, Solak Y, *et al.*. Uric acid in metabolic syndrome: From an innocent bystander to a central player. *Eur J Intern Med* 2016;29:3-8. 10.1016/j.ejim.2015.11.026 [PMCID: PMC4826346] [PubMed: 26703429] [CrossRef: 10.1016/j.ejim.2015.11.026]

112. Wu T, Giovannucci E, Pischon T, *et al.*. Fructose, glycemic load, and quantity and quality of carbohydrate in relation to plasma C-peptide concentrations in US women. *Am J Clin Nutr* 2004;80:1043-9. 10.1093/ajcn/80.4.1043 [PubMed: 15447918] [CrossRef: 10.1093/ajcn/80.4.1043]

113. Quiñ ones Galvan A, Natali A, Baldi S, *et al.*. Effect of insulin on uric acid excretion in humans. *Am J Physiol* 1995;268:E1-5. [PubMed: 7840165]

114. Richette P, Bardin T. Gout. *Lancet* 2010;375:318-28. 10.1016/S0140-6736(09)60883-7 [PubMed: 19692116] [CrossRef: 10.1016/S0140-6736(09)60883-7]

115. Nakagawa T, Tuttle KR, Short RA, Johnson RJ. Hypothesis: fructose-induced hyperuricemia as a causal mechanism for the epidemic of the metabolic syndrome. *Nat Clin Pract Nephrol* 2005;1:80-6. 10.1038/ncpneph0019 [PubMed: 16932373] [CrossRef: 10.1038/ncpneph0019]

116. Cicero AFG, Fogacci F, Desideri G, *et al.*. Arterial Stiffness, Sugar-Sweetened Beverages and Fruits Intake in a Rural Population Sample: Data from the Brisighella Heart Study. *Nutrients* 2019;11:2674. 10.3390/nu11112674 [PMCID: PMC6893603] [PubMed: 31694231] [CrossRef: 10.3390/nu11112674]

117. Caliceti C, Calabria D, Roda A, Cicero AFG. Fructose Intake, Serum Uric Acid, and Cardiometabolic Disorders: A Critical Review. *Nutrients* 2017;9:395. 10.3390/nu9040395 [PMCID: PMC5409734] [PubMed: 28420204] [CrossRef: 10.3390/nu9040395]

118. Johnson RJ, Sanchez-Lozada LG, Nakagawa T. The effect of fructose on renal biology and disease. *J Am Soc Nephrol* 2010;21:2036-9. 10.1681/ASN.2010050506 [PubMed: 21115612] [CrossRef: 10.1681/ASN.2010050506]

119. World Cancer Research Fund International. Diet, nutrition, physical activity and cancer: A global perspective. A summary of the Third Expert Report. 2018. https://www.wcrf.org/dietandcancer/a-summary-of-the-third-expert- report/.

120. Augustin LSA, Kendall CWC, Jenkins DJA, *et al.*. Glycemic index, glycemic load and glycemic response: An International Scientific Consensus Summit from the International Carbohydrate Quality Consortium (ICQC). *Nutr Metab Cardiovasc Dis* 2015;25:795-815. 10.1016/j.numecd.2015.05.005 [PubMed: 26160327] [CrossRef: 10.1016/j.numecd.2015.05.005]

121. Avgerinos KI, Spyrou N, Mantzoros CS, Dalamaga M. Obesity and cancer risk: Emerging biological mechanisms and perspec-

tives. *Metabolism* 2019;92:121-35. 10.1016/j.metabol.2018.11.001 [PubMed: 30445141] [CrossRef: 10.1016/j.metabol.2018.11.001]

122. Todoric J, Di Caro G, Reibe S, *et al.*. Fructose stimulated de novo lipogenesis is promoted by inflammation. *Nat Metab* 2020;2:1034-45. 10.1038/s42255-020-0261-2 [PMCID: PMC8018782] [PubMed: 32839596] [CrossRef: 10.1038/s42255-020-0261-2]
123. Do MH, Lee E, Oh MJ, Kim Y, Park HY. High-Glucose or -Fructose Diet Cause Changes of the Gut Microbiota and Metabolic Disorders in Mice without Body Weight Change. *Nutrients* 2018;10:761. 10.3390/nu10060761 [PMCID: PMC6024874] [PubMed: 29899272] [CrossRef: 10.3390/nu10060761]
124. Goncalves MD, Lu C, Tutnauer J, *et al.*. High-fructose corn syrup enhances intestinal tumor growth in mice. *Science* 2019;363:1345-9. 10.1126/science.aat8515 [PMCID: PMC6487857] [PubMed: 30898933] [CrossRef: 10.1126/science.aat8515]
125. Harrell CS, Burgado J, Kelly SD, Johnson ZP, Neigh GN. High-fructose diet during periadolescent development increases depressive-like behavior and remodels the hypothalamic transcriptome in male rats. *Psychoneuroendocrinology* 2015;62:252-64. 10.1016/j.psyneuen.2015.08.025 [PMCID: PMC4637272] [PubMed: 26356038] [CrossRef: 10.1016/j.psyneuen.2015.08.025]
126. Avena NM, Rada P, Hoebel BG. Evidence for sugar addiction: behavioral and neurochemical effects of intermittent, excessive sugar intake. *Neurosci Biobehav Rev* 2008;32:20-39. 10.1016/j.neubiorev.2007.04.019 [PMCID: PMC2235907] [PubMed: 17617461] [CrossRef: 10.1016/j.neubiorev.2007.04.019]
127. Tucker KL. Dietary intake and bone status with aging. *Curr Pharm Des* 2003;9:2687-704. 10.2174/1381612033453613 [PubMed: 14529541] [CrossRef: 10.2174/1381612033453613]
128. Alswat KA. Gender Disparities in Osteoporosis. *J Clin Med Res* 2017;9:382-7. 10.14740/jocmr2970w [PMCID: PMC5380170] [PubMed: 28392857] [CrossRef: 10.14740/jocmr2970w]
129. Lustig RH, Schmidt LA, Brindis CD. Public health: The toxic truth about sugar. *Nature* 2012;482:27-9. 10.1038/482027a [PubMed: 22297952] [CrossRef: 10.1038/482027a]
130. Tahmassebi JF, Duggal MS, Malik-Kotru G, Curzon ME. Soft drinks and dental health: a review of the current literature. *J Dent* 2006;34:2-11. 10.1016/j.jdent.2004.11.006 [PubMed: 16157439] [CrossRef: 10.1016/j.jdent.2004.11.006]
131. Moynihan PJ, Kelly SA. Effect on caries of restricting sugars intake: systematic review to inform WHO guidelines. *J Dent Res* 2014;93:8-18. 10.1177/0022034513508954 [PMCID: PMC3872848] [PubMed: 24323509] [CrossRef: 10.1177/0022034513508954]
132. Kouba J. Impact of Environment, Ethnicity, and Culture on Nutrition and Health. In: Touger-Decker R, Sirois DA, Mobley CC, eds. *Nutrition and Oral Medicine*. Humana Press, 2005:45-60 10.1385/1-59259-831-5:045. [CrossRef: 10.1385/1-59259-831-5:045]
133. Pestoni G, Krieger JP, Sych JM, Faeh D, Rohrmann S. Cultural Differences in Diet and Determinants of Diet Quality in Switzerland: Results from the National Nutrition Survey menuCH. *Nutrients* 2019;11:126. 10.3390/nu11010126 [PMCID: PMC6357532] [PubMed: 30634520] [CrossRef: 10.3390/nu11010126]
134. Dhingra R, Sullivan L, Jacques PF, *et al.*. Soft drink consumption and risk of developing cardiometabolic risk factors and the metabolic syndrome in middle-aged adults in the community. *Circulation* 2007;116:480-8. 10.1161/CIRCULATIONAHA.107.689935 [PubMed: 17646581] [CrossRef: 10.1161/CIRCULATIONAHA.107.689935]
135. Gardener H, Rundek T, Markert M, Wright CB, Elkind MS, Sacco RL. Diet soft drink consumption is associated with an increased risk of vascular events in the Northern Manhattan Study. *J Gen Intern Med* 2012;27:1120-6. 10.1007/s11606-011-1968-2 [PMCID: PMC3514985] [PubMed: 22282311] [CrossRef: 10.1007/s11606-011-1968-2]
136. Eshak ES, Iso H, Kokubo Y, *et al.*. Soft drink intake in relation to incident ischemic heart disease, stroke, and stroke subtypes in Japanese men and women: the Japan Public Health Centre-based study cohort I. *Am J Clin Nutr* 2012;96:1390-7. 10.3945/ajcn.112.037903 [PubMed: 23076619] [CrossRef: 10.3945/ajcn.112.037903]
137. Fung TT, Malik V, Rexrode KM, Manson JE, Willett WC, Hu FB. Sweetened beverage consumption and risk of coronary heart disease in women. *Am J Clin Nutr* 2009;89:1037-42. 10.3945/ajcn.2008.27140 [PMCID: PMC2667454] [PubMed: 19211821] [CrossRef: 10.3945/ajcn.2008.27140]
138. Pereira MA, Kartashov AI, Ebbeling CB, *et al.*. Fast-food habits, weight gain, and insulin resistance (the CARDIA study): 15-year prospective analysis. *Lancet* 2005;365:36-42. 10.1016/S0140-6736(04)17663-0 [PubMed: 15639678] [CrossRef: 10.1016/S0140-6736(04)17663-0]
139. Jaeschke R, Guyatt GH, Dellinger P, *et al.*GRADE Working Group . Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 2008;337:a744. 10.1136/bmj.a744 [PubMed: 18669566] [CrossRef: 10.1136/bmj.a744]
140. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans . Some chemicals present in industrial and consumer products, food and drinking-water. *IARC Monogr Eval Carcinog Risks Hum* 2013;101:9-549. [PMCID: PMC4939416] [PubMed: 24772663]
141. Smith TJ, Wolfson JA, Jiao D, *et al.*. Caramel color in soft drinks and exposure to 4-methylimidazole: a quantitative risk assessment. *PLoS One* 2015;10:e0118138. 10.1371/journal.pone.0118138 [PMCID: PMC4333292] [PubMed: 25693062] [CrossRef: 10.1371/journal.pone.0118138]
142. Nardin T, Barnaba C, Abballe F, Trenti G, Malacarne M, Larcher R. Fast analysis of quaternary ammonium pesticides in food and beverages using cation-exchange chromatography coupled with isotope-dilution high- resolution mass spectrometry. *J Sep Sci* 2017;40:3928-37. 10.1002/jssc.201700579 [PubMed: 28779575] [CrossRef: 10.1002/jssc.201700579]
143. Albero B, Sánchez-Brunete C, Tadeo JL. Determination of organophosphorus pesticides in fruit juices by matrix solid-phase dispersion and gas chromatography. *J Agric Food Chem* 2003;51:6915-21. 10.1021/jf030414m [PubMed: 14611147] [CrossRef: 10.1021/jf030414m]
144. Additives E. Scientific Opinion on the re-evaluation of aspartame (E 951) as a food additive. *EFSA J* 2013;11:3495.

145. Jayalath VH, de Souza RJ, Ha V, *et al.*. Sugar-sweetened beverage consumption and incident hypertension: a systematic review and meta-analysis of prospective cohorts. *Am J Clin Nutr* 2015;102:914-21. 10.3945/ajcn.115.107243 [PubMed: 26269365] [CrossRef: 10.3945/ajcn.115.107243]

146. Joh HK, Lee DH, Hur J, *et al.*. Simple Sugar and Sugar-Sweetened Beverage Intake During Adolescence and Risk of Colorectal Cancer Precursors. *Gastroenterology* 2021;161:128-142.e20. 10.1053/j.gastro.2021.03.028 [PMCID: PMC8238879] [PubMed: 33753105] [CrossRef: 10.1053/j.gastro.2021.03.028]

147. Malik VS, Li Y, Pan A, *et al.*. Long-Term Consumption of Sugar-Sweetened and Artificially Sweetened Beverages and Risk of Mortality in US Adults. *Circulation* 2019;139:2113-25. 10.1161/CIRCULATIONAHA.118.037401 [PMCID: PMC6488380] [PubMed: 30882235] [CrossRef: 10.1161/CIRCULATIONAHA.118.037401]

148. Gao M, Jebb SA, Aveyard P, *et al.*. Associations Between Dietary Patterns and Incident Type 2 Diabetes: Prospective Cohort Study of 120,343 UK Biobank Participants. *Diabetes Care* 2022;45:1315-25. 10.2337/dc21-2258 [PubMed: 35299247] [CrossRef: 10.2337/dc21-2258]

Kannabinoideek *in vitro* ugalketaren eraginkortasuna hobetu dezakete

Antzutasuna detektatzeko eta tratatzeko metodologia berriak aurkeztu ditu ikerketa batek, emakumeen sistema kannabinoide fisiologikoaren gabeziek ugalkortasunean eragiten dituzten arazoak identifikatu ondoren. Halaber, ondorioztatu dute kalamu landarearen osagai batzuek aukera eman dezaketela in vitro ugalketaren hormonazio-tratamenduak ekiditeko, zeinak oso gogorrak baitira emakumeentzat.

Animaliek kalamuaren antzeko osagaiak dituen sistema fisiologiko bat dute. Hau da, barne-sistema kannabinoide bat, zelulen arteko komunikaziorako sistema moduan funtzionatzen duena. Organismoko prozesu fisiologiko garrantzitsu asko kontrolatzen ditu. Besteak beste, ugalkortasuna. Obuluen eta espermatozoideen sorreran ez ezik, ernalketan, enbrioia umetokian ezartzeko prozesuan, plazentaren sorreran eta erditzean bertan dago inplikatuta.

Azken urteotako ikerketek erakutsi dute sistema endokannabinoidean gabeziaren bat izateak akats larriak eragin ditzakeela ugalkortasunean. Orain arte ikusi izan dute obulazioa atzeratzen duela, baina, orain, ikerketa batek erakutsi du obarioak hondatu, eta folikuluen zein obozitoen garapena ere eragozten duela. Beraz, uste dute sistema endokannabinoidearen gabeziak antzutasuna detektatzeko biomarkatzaile gisa erabil daitezkeela. Hartara, interesgarria izan liteke sistema endokannabinoidearen gabeziak detektatzeko probak antzutasuna detektatzeko ohiko probetan txertatzea. Baliteke, beraz, antzutasuna tratatzeko terapia berriak ere sortzea.

Baina kannabinoide exogenoen aukerak ere aztertu dituzte, eta ondorioztatu dute kalamu landarearen osagai batzuk baliagarriak izan daitezkeela *in vitro* ugalkortasun-tratamenduetan.

Kalamuaren kontsumoak ugalketa galaraz dezake

Kannabinoide exogenoen kontsumo jarraituak kalteak sortzen ditu ugalketa-aparatuan. Kannabinoide exogenoek endogenoekin dituzten antzekotasunak direla medio, organismoko sistema endokannabinoideak erregulatzen dituen ugalketa-prozesuak asaldatzen ditu kontsumo horrek: arriskuan jartzen ditu enbrioiaren hasierako garapena eta biziraupena, eta gerta liteke haurdunaldia bertan behera gelditzea.

Alabaina, erabilera terapeutikoa ere izan dezakete, pazienteek beren sistema endokannabinoidean akatsak dituzten kasuetan. Zehazki, kalamuaren osagai psikoaktibo nagusiarekin egindako ikerketek frogatu dute alternatiba egokia izan daitekeela *in vitro* ugalketa-tekniken laguntza behar duten emakumeentzat.

Obozitoak *in vitro* heltzeko teknikan erabil daiteke. Oraindik gutxi erabiltzen da teknika hori *in vitro* ernalketan, baina interes handikoa da, egun erabiltzen diren tratamendu hormonal ohikoek albo-ondorio larriak eragin baititzakete: obarioen hiperestimulazio-sindromea eta bestelako arazoak. Beraz, emakume denentzat ez da gomendagarria halako tratamenduak hartzea. Obozitoak *in vitro* heltzeko teknikak, ordea, ez du halako tratamendu hormonalik behar, emakumeari obozitoak heldu gabe erauzi eta organismotik kanpo heltzen baitira. Ikerketa honek frogatu du kannabinoide exogenoek laguntzen dutela pazienteari erauzitako obozitoak laborategian heltzen.

Ondorio garrantzitsuak dira, emakume askoren osasunak ez baitu ahalbidetzen tratamendu hormonalak hartzea. Europan, urtetik urtera handituz doa *in vitro* ugalketa behar duten pertsonen kopurua; batetik, ugalketa-adina handitu delako, eta, bestetik, gero eta ohikoagoak direlako seme-alabak izan nahi dituzten guraso bakarreko familiak eta bikote homosexualak. Beraz, emakume askorentzat izan liteke alternatiba egokia.

InpAkta 00 (2023) 99-106
https://doi.org/10.1387/ekaia.22702

Sistema endokannabinoideak emearen ugal-fisiologia modulatzen du

Lide Totorikaguena Iturriaga*, Estibaliz Olabarrieta López, Naiara Agirregoitia Marcos, Ekaitz Agirregoitia Marcos
Fisiologia saila, Medikuntza eta Erizaintza Fakultatea (UPV/EHU)

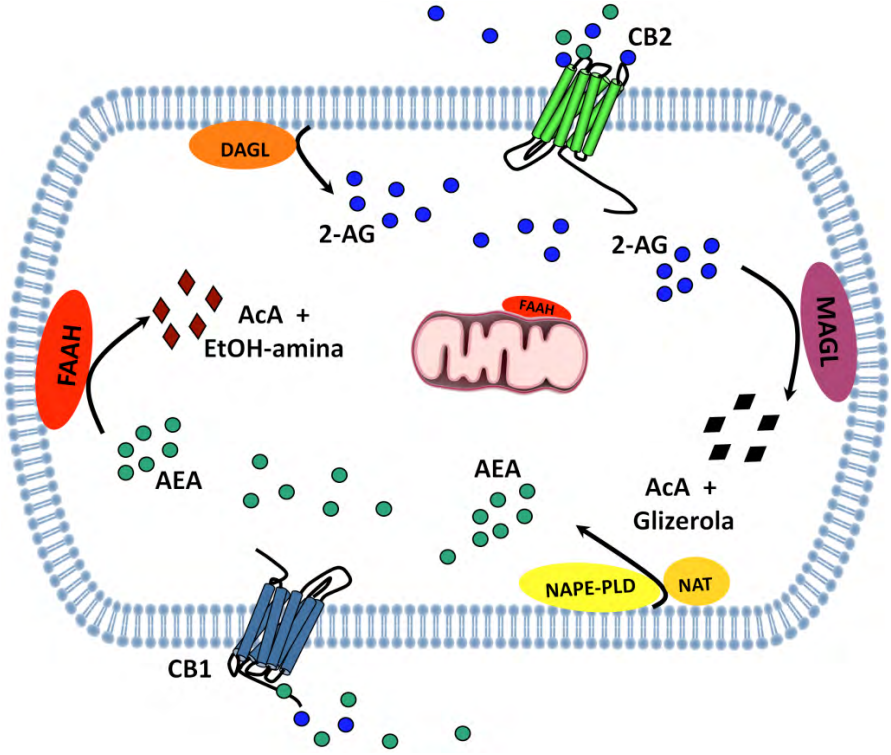
LABURPENA: Sistema kannabinideak ugal-sistemaren prozesu desberdinetan parte hartzen du, eta jakina da kannabinoide exogenoen kontsumo kroniko sistemikoa kaltegarria dela ugalketa-prozesuetarako; sistema endokannabinoidean izan daitezkeen gabeziak, halaber, ez dira mesedegarriak. Frogatu da sistema endokannabinoidearen edozein osagairen aktibazioak edo eraldaketek eragina izan dezaketela ugalketaren fase desberdinen erregulazioan eta bizitza berri baten hasiera arriskuan jar dezaketela. Horregatik, kannabinoideen erabilera lagungarria izan liteke, antzutasunaren diagnostikoan biomarkatzaile posible gisa edota praktika klinikoan tratatzeko itu terapeutiko posible bezala.

1. Sarrera

Kannabinoideak kalamu landareak (*Cannabis sativa L.*) sintetizatutako metabolito sekundarioak dira, baina *Cannabis sativa* landarea beste konposatu kimiko ugarik osatzen dute. Identifikatutako 483 konposatuetatik 60 kannabinoide baino gehiago ageri dira kannabisaren landarean. 1964. urtetik aurrera nagusitu zen kannabinoideen inguruko ikerketa, landare honen Δ9-tetrahidrokannabinola (THC) [1] osagai psikoaktibo ugariena aurkitu zenean. Izan ere, THC da efektu bioaktiboen arduradun nagusia eta kannabinoide ugariena.

Bitxia badirudi ere, barne-sistema kannabinoide bat (sistema endokannabinoidea) deskribatu zen animalietan [2]. Sistema hori honako hauek osatzen dute: kannabinoide-hartzaileek (CB1 eta CB2), barne-estekatzaileek (anandamida (AEA) eta 2-arakidonoilglizerol (2-AG) endokannabinoideak) eta sintesi- [N-arakidonoilfosfatidiletanolamina (NAPE-PLD), N-aziltransferasa (NAT) eta diazilglizerol lipasa (DAGL)] eta degradazio-entzimek [gantz-azidoen amida hidrolasa (FAAH) eta monoazilglizerol lipasa (MAGL)] [2] (1. irudia).

* **Corresponding author / Harremanetan jartzeko:** Lide Totorikaguena Iturriaga. Fisiologia saila, Medikuntza eta Erizaintza Fakultatea (UPV/EHU), Leioa (Bizkaia). e-mail: lide.totorikaguena@ehu.eus. https://orcid.org/0000-0002-8315-4608 **Contributed by** Ugo Mayor **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Totorikaguena Iturriaga, Lide *et al.* «Sistema endokannabinoideak emearen ugal-fisiologia modulatzen du», *Ekaia* | (2022)42 (https://doi.org/10.1387/ekaia.22702)



1. irudia: Barne-sistema kannabinoidearen osagai nagusiak: kannabinoide-hartzaileak (CB1 eta CB2), barne-estekatzaileak (AEA eta 2-AG endokannabinoideak) eta sintesi- (NAPE-PLD, NAT eta DAGL) eta degradazio-entzimak (FAAH eta MAGL).

Kannabinoideen inguruko ikerketek ondorioztatu dute sistema kannabinoidearen funtzio nagusietarikoa neuromodulazioa dela, CB1 hartzailea G proteinei loturiko hartzailerik ugariena baita nerbio-sistema zentralean. Hala ere, CB1 eta CB2 hartzaileen presentzia eta, oro har, sistema endokannabinoidearen modulazioa eta ekintza-mekanismoak, nerbio-sistema zentralean agertzeaz gain, ehun periferikoetan ere deskribatu dira. Sistema kannabinoidea inplikatua dagoen prozesu horietako bat ugalkortasuna da. Sistema hori gametogenesian, ernalketan, enbrioiaren ezarpenean, plazentazioan, haurdunaldian eta erditzean identifikatu da [3, 4] eta sistema kannabinoidean parte hartzen duen makinaria guztia ugal-aparatuko organo, ehun eta zeluletan dagoela frogatu da [5, 6].

Horregatik, sistema endokannabinoidea biomarkatzaile posible gisa proposatu da praktika klinikoan ugalketa potentziala aurreikusteko [7]. Horretarako, animalia-eredu desberdinekin egindako esperimentuek erakutsitakoa berrikusi dugu, bai eta ugalketa-patologien inguruan egin den ikerketa ere.

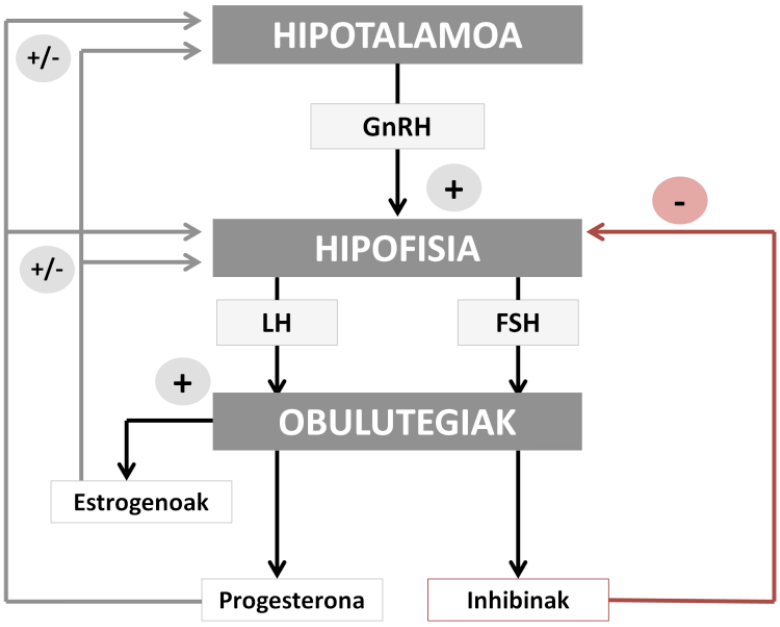
2. Sistema kannabinoidea emearen ugalkortasunean

Emeari dagokionez, sistema endokannabinoidea deskribatu da hainbat ugaztun espezietako hipotalamo-hipofisi-obulutegiaren ardatzean [5], likido folikularrean, obulutegietan, obozitoetan eta pikor-geruzako zeluletan [8, 9, 10, 11, 12, 13, 14, 15].

2.1. *Sistema kannabinoidea hipotalamo-hipofisi-obario (HPO) ardatzean*

Sistema endokannabinoidea hainbat estimulu fisiologikok kontrolatzen dute: besteak beste, hormona-mailak. Sistema endokannabinoidea GnRH bezalako hormonen ekoizpenaz arduratzen den hipotalamoaren guneetan antzeman da, zeinek hipotalamo-hipofisi-obario (HPO) ardatzaren bidez jarduten duten ugalketa-prozesuen hainbat alderdi kontrolatzeko [16] (2. irudia). Oro har, sistema endokannabinoidearen efektoreek eragin handia dute ugalkortasunean eta funtzio endokrinoan, karraskariekin, primateekin eta gizakiekin

egindako ikerketek frogatu bezala [9, 17]. Kannabisak ugalketaren fisiologiaren hainbat alderditan duen eragina azaldu lezake horrek, besteak beste, hormonen askatzean HPO ardatzean [18]. Gonadotropinek, progesteronak eta estrogenoak, esaterako, AEA mailak [19] eta FAAH entzimaren adierazpena erregulatzen dituzte hilekoan zehar [20].



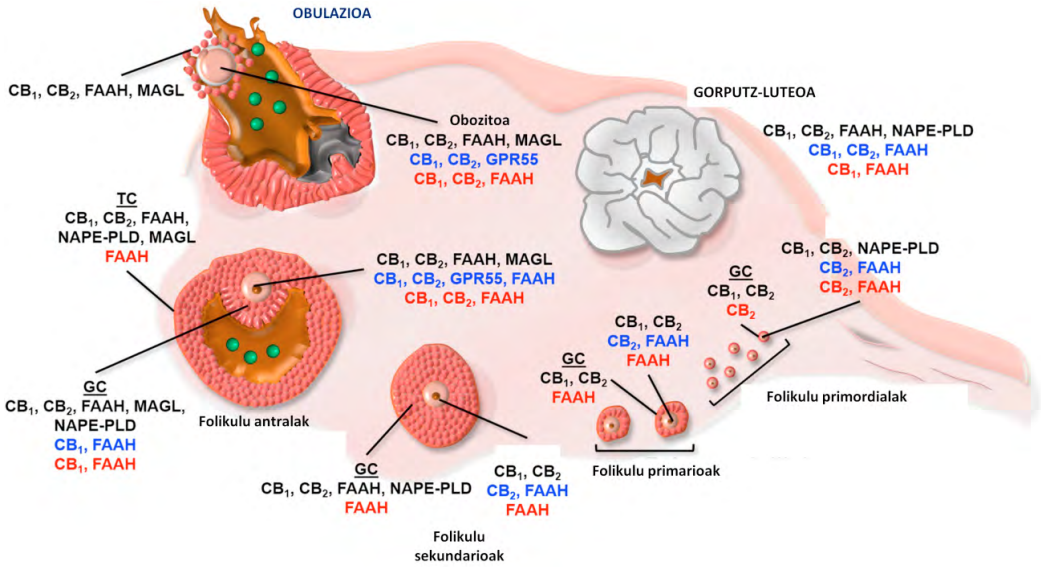
2. irudia: Hipotalamo-hipofisi-obario ardatzaren eskema orokorra. Hipotalamoak GnRH jariatzen du eta honek hipofisia estimulatzen du. Horrela, hipofisiak FSH eta LH jariatzen ditu odol-fluxura obulutegietaraino eta obulutegietan; ondorioz, folikuluak garatzen hasten dira estradiola jariatuz. LHk obulazioa eragiten du, bai eta luteo-gorputz-luteoa eratzea ere, progesterona eta estrogenoak jariatuz.

Sistema endokannabinoidea kanpotik modulatzen denean, kannabinoideek HPO ardatzaren erregulazioa asaldatzen dute, ugaltze-sistema kaltetuz (adibidez, obulaziorik gabeko hilekoak izatea) [16]. Gainera, GnRH hormonaren jariaketa murrizten dute [16] eta hormona horrek gonadotropinen (FSHa eta LHa) askapena ekiditen du, gonaden funtzioa erabat kaltetuz [5]. Era berean, sistema endokannabinoidearen adierazpenean egon litezkeen kalteak antzutasunarekin erlazionatuta daude [5], jakina baita kannabinoide-hartzaileen gabeziak HPO ardatzeko hormonen askapena inhibitzen duela: besteak beste, GnRHa, FSHa, 17-β-estradiola [21] eta LHarena [22, 23].

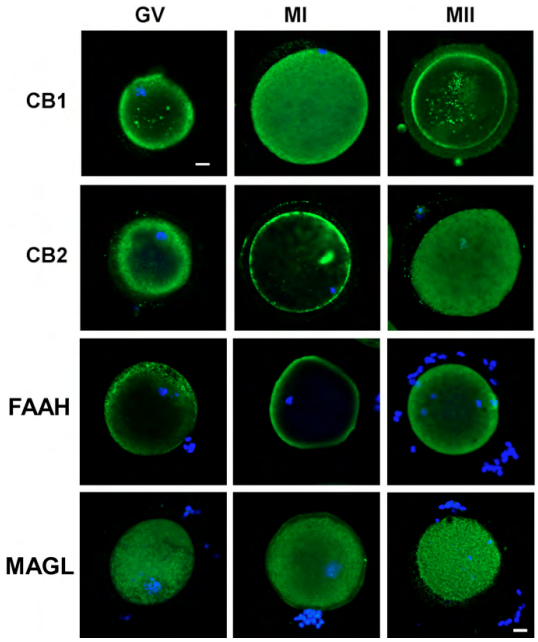
Hitz gutxitan, sistema endokannabinoidea eta HPO ardatza erlazionatuta daudela baieztatu da, nahiz eta hori kontrolatzen duten mekanismoak oraindik ez diren guztiz ezagutzen [5].

2.2. *Sistema kannabinoidea obulutegietan: folikulugenesian, obozitoen garapenean eta heltze-prozesuan*

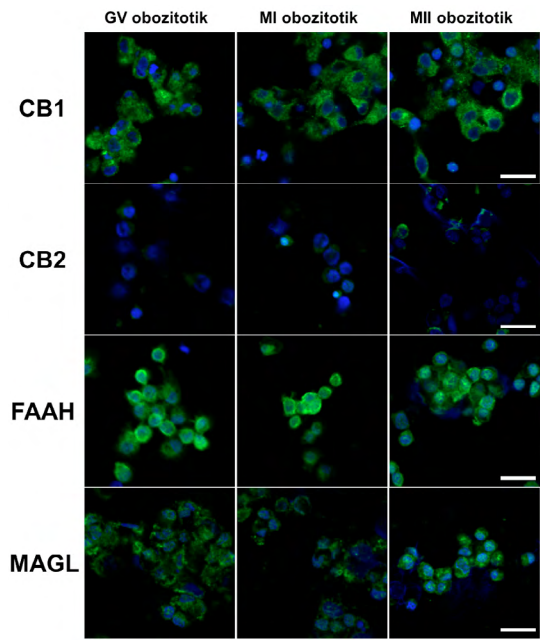
Kannabinoideen eta ugalketa-sistemaren arteko lotura, lehenen aldiz, likido folikularrean endokannabinoide bat, AEA, aurkitzeari esker egin zen [8]. Frogatu zuten bere kontzentrazioa aldakorra zela obulazioan eta haurdunaldi goiztiarrean zehar [18]. Horrez gain, bai AEA bai endokannabinoide hori metabolizatzen duten entzimak gizakiaren obulutegian lokalizatu ziren [9], eta, ondoren, sistema endokannabinoide osoa deskribatu zen giza [9] eta arratoien [10] obarioetan eta folikuluetan. Ikerketa horiei guztiei esker, kannabinoide-hartzaileen eta degradazio-entzimen lokalizazioa deskribatu zen (3. irudia). Beraz, horrek guztiak iradokitzen du endokannabinoideen seinaleztapenak folikulugenesia erregulatzen lagundu lezakeela [9, 10].



3. irudia: Sistema endokannabinoidearen osagaien banaketa ugaztunen obulutegian zehar. Folikulu bakoitza obozitoak eta inguruan dituen pikor-geruzako zelulek (GC) osatzen dute; teka-zelulak (TC) geroago agertzen dira folikuluen garapenean zehar. Gonadotropinak (FSH, LH) beharrezkoak dira folikuluen garapena eta obulazioa sustatzeko. Sistema endokannabinoidearen osagaien adierazpena desberdin modulatzen da folikulogenesian zehar, eta espezieen artean (beltzez, gizakiak; urdinez, karraskariak; gorriz, beste ugaztun batzuk) (*Cecconi eta lank., 2020*-tik moldatua) [6].



4. irudia: Kannabinoide-hartzaileen (CB1 eta CB2) eta degradazio-entzimen (FAAH eta MAGL) kokapen immunozitokimikoa giza obozitoen heltze-prozesuaren faseetan. CB1 eta CB2 kannabinoide-hartzaileen eta FAAH eta MAGL degradazio-entzimen banaketak berdez ageri dira besikula germinalean (GV), obozitoa meiosiko profase I-ean dagoenean (MI) eta obozitoa meiosiko bigarren metafasean dagoenean (MII). Hoechst DNA-markatzailea, urdinez. Erreferentzia barra: 20 µm [11].



5. irudia: Kannabinoide-hartzaileen (CB1 eta CB2) eta degradazio-entzimen (FAAH eta MAGL) kokapen immunozitokimikoa pikor-geruzako zeluletan giza obozitoen heltze-prozesuaren fase desberdinetan. CB1, CB2 kannabinoide-hartzaileen eta FAAH eta MAGL degradazio-entzimen banaketak berdez ageri dira besikula germinaletik eratorritako (GV) pikor-geruzako zeluletan, obozitoa meiosiko profase I-ean dagoen (MI) eta obozitoa meiosiko bigarren metafasean dagoen (MII) pikor-geruzako zeluletan. Hoechst DNA-markatzailea, urdinez. Erreferentzia barra: 10 µm [14].

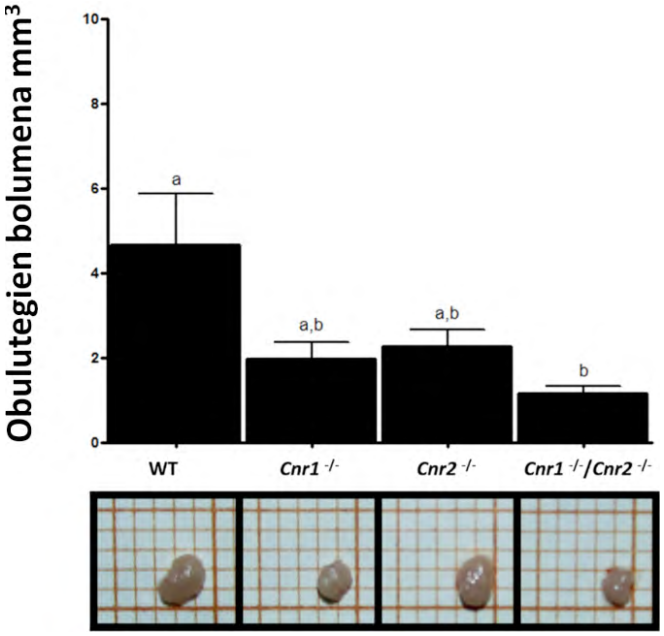
Duela urte gutxi, gure ikerketa taldeak sistema kannabinoidea gizakien [11], behien [12] eta saguen obozitoetan adierazten dela baieztatu zuen [13]. Zehazki, obozitoen meiosiaren berraktibazioan CB1 eta CB2 kannabinoide-hartzaileak identifikatu genituen, bai RNA mezulari-mailan bai proteina-mailan ere. Gainera, giza obozitoetan kannabinoideentzako degradazio-entzimak, FAAH eta MAGL, espresatzen zirela deskribatu genuen (4. irudia). Are gehiago, obozitoaren heltze-prozesuan CB1, FAAH eta MAGL proteinen denborazko kokapena aldatuz joaten dela erakutsi genuen [11, 15].

Beste alde batetik, giza pikor-geruzako zeluletan CB1 eta CB2 kannabinoide-hartzaileak eta endokannabinoideen andeatzea katalizatzen dituzten bi entzimak, FAAH eta MAGLa, lokalizatu genituen obozitoaren heltze-prozesuan zehar [14] (5. irudia). Aurkikuntza hori interesgarria izan zen; izan ere, obulazioan, pikor-geruzako kumuluko zelulak-obozito konplexua (ingelesez, cumulus-oocyte complex, COC) askatzen da, eta obozitoaren heltze-prozesuan oso garrantzitsua da obozitoaren eta pikor-geruzako zelulen arteko bi norabideko komunikazioa. Hau da, pikor-geruzako zelulek ere obozitoaren heltze-prozesuan parte hartzen duten mekanismoak erregulatzen dituzte [24].

3. Sistema endokannabinoidearen gabeziaren eragina obulutegiaren morfologian, folikulugenesian eta obozitoen heltze-prozesuan

Ondorioztatu da kannabinoide exogenoen kontsumo kroniko sistemikoa kaltegarria dela ugalketa-prozesuetarako, kannabinoideak kanpotik hartzean sistema endokannabinoidearen seinaleztapena asaldatzen delako. Esaterako, tximinoekin egindako ikerketa batean ikusi zen THCaren kontsumo kronikoa obulazioaren atzerapenarekin erlazionatuta dagoela [25].

Beraz, jakina da sistema endokannabinoidearen edozein osagairen aktibazioak edo eraldaketek eragina izan dezaketela ugalketaren fisiologian eta, hortaz, sistema endokannabinoidean egon daitezkeen gabeziak ere kaltegarriak dira. Saguekin egindako azterketa batzuek iradokitzen dutenez, obozitoetan kannabinoide-hartzaileen gabezia egoteak, ernalketaren ondorengo enbrioiaren garapenean akatsak eragiten ditu eta [13] CB1 hartzailearentzako *knockout* genotipoko saguen % 40 inguruk haurdunaldiaren galera erakusten du [26].



6. irudia: Kannabinoide-hartzaileen gabeziaren eragina saguen obulutegien bolumenean. WT: genotipo basatidun saguak. Cnr1$^{-/-}$: CB1 hartzailea falta zaien *knockout* saguak. Cnr2$^{-/-}$: CB2 hartzailea falta zaien *knockout* saguak. Cnr1$^{-/-}$/Cnr2$^{-/-}$: CB1 eta CB2 hartzaileak falta zaizkien *knockout* saguak [26].

Alta, orain arte ez zen ezagutzen zer gertatzen zen obarioen morfologian, folikulugenesian eta obozitoen heltze-prozesuan kannabinoideen seinaleztapena ezabatzen zenean, eta hori izan da gure esperimentuek ekarri duten berrikuntza. Zentzu horretan, gure emaitzek erakutsi zuten CB1 hartzailearen gabeziak efektu kaltegarriak zituela obarioaren morfologian (kontrolarekin alderatuz, azalera eta bolumen txikiagoko obulutegiak dituzte) (6. irudia), folikulugenesian (folikulu gutxiago dituzte kontrolarekin alderatuz eta eCG tratamenduaren aurrean ez dute ondo erantzuten) eta obulatzen duten obozitoen kalitatean (obozito hedatuen kopuru baxuagoa erakusten baitute) eta meiosiaren progresioan (obozitoaren heltze-proesuan sinkronizazio falta nabari da, obozito helduak beranduago eskuratzen baitira). Are gehiago, CB2 hartzailearen galerak aipatu berri ditugun pro-

zesu horietan akats hain esanguratsuak sortzen ez dituen arren, kannabinoide-hartzaile biak ez egoteak areagotu egiten zuen kaltea, batez ere obulutegien tamaina txikiagoak eta folikulu kopuru gutxiago aurkituz; CB2 hartzaileak ere garrantzia zuela erakutsiz. Ez hori bakarrik, eCG hormonaren eraginkortasuna ere kaltetua ikusi da kannabinoide-hartzaile biak falta zaizkien animalietan [27]. Horregatik, badirudi kannabinoide-hartzaileen faltak obulutegien bizitza funtzionala eta kalitatea okertzen dituela.

Gainera, kontuan izan behar da antzutasuna obulutegien folikuluen kantitatearen eta kalitatearen beherakadaren ondorioa dela [28] eta obozitoen garapen-gaitasuna folikuluen tamainak eta kalitateak baldintzatua egon daitekeela [29].

4. Ondorioak

Hortaz, bai kannabinoideen gehiegizko esposizioa zein sistema endokannabinoidearen osagairen baten gabezia kaltegarriak izan daitezkeenez ugal-prozesuetan, ondoriozta genezake sistema endokannabinoideak prozesu horiek modulatzen dituela eta garrantzitsua dela barne kannabinoideen kontzentrazioa ondo erregulatua egotea. Horregatik, kannabinoideen erabilera aztertzen hasia da, antzutasunaren diagnostikoan biomarkatzaile posible gisa edota praktika klinikoan tratatzeko itu terapeutiko posible gisa. Esaterako, gure ikerketa-taldeak lerro interesgarri bat zabaldu du kannabinoideen erabilera terapeutikoa aztertuz obozitoen *in vitro* heltzeko (IVM) medioen eraginkortasuna hobetzeko. Posibletzat jotzen zen endokannabinoideen seinaleztapenak folikulugenesia eta obozitoen heltze-prozesuak modula zitzakeela [8, 9], eta saguetan eta behietan egindako ikerketek frogatu dute lagunduriko ugalketa-tekniken eraginkortasuna hobetu daitekeela, teknika horietan erabiltzen diren hazkuntza-medioak kanabinoideekin aberastuz gero [12, 13, 27, 30].

5. Bibliografia

[1] Mechoulam, R., eta Gaoni, Y. 1967. «The absolute configuration of delta-1-tetrahydrocannabinol, the major active constituent of hashish». *Tetrahedron Lettera*. **8**, 1109-1111.

[2] Correa, F., Wolfson, M. L., Valchi, P., Aisemberg, J., eta Franchi, A. M. 2016. «Endocannabinoid system and pregnancy». *Reproduction*. **152**, R191-R200.

[3] Battista, N., Rapino, C., Di Tommaso, M., Bari, M., Pasquariello, N. eta Maccarrone, M. 2008. «Regulation of male fertility by the endocannabinoid system». *Mol Cell Endocrinol*. **286**, S17-S23

[4] Maccarrone, M. 2009. «Endocannabinoids: friends and foes of reproduction». *Progress in lipid research*. **48**, 344-354.

[5] Brents, L. K. 2016. «Marijuana, the endocannabinoid system and the female reproductive system». *Yale Journal of Biology and Medicine*. **89**, 175-191.

[6] Cecconi, S., Rapino, C., Di Nisio, V., Rossi, G. eta Maccarrone, M. 2020. «The (endo)cannabinoid signaling in female reproduction: What are the latest advances?».*Prog Lipid Res*. 77, 101019

[7] Sun, X. eta Dey, SK. 2012. «Endocannabinoid signaling in female reproduction». *ACS Chem Neurosci*. **3**, 349-55.

[8] Schuel, H., Burkman, L. J., Lippes, J., Crickard, K., Mahony, M. C., Giuffrida, A. eta Makriyannis, A. 2002. «Evidence that anandamidesignaling regulates human sperm functions required for fertilization». *Molecular Reproduction and Development*. **63**, 376-387.

[9] El-Talatini, M. R., Taylor, A. H., Elson, J. C., Brown, L., Davidson, A. C. eta Konje, J. C. 2009. «Localisation and function of the endocannabinoid system in the human ovary». *PLOS One*. **4**, e4579.

[10] Bagavandoss, P. eta Grimshaw, S. 2010. «Temporal and spatial distribution of the cannabinoid receptors (CB 1, CB2) and fatty acid amide hydroxylase in the rat ovary». *Anatomical Record*. **293**, 1425-1432.

[11] Peralta, L., Agirregoitia, E., Mendoza, R., Expósito, A., Casis, L., Matorras, R. eta Agirregoitia, N. 2011. «Expression and localization of cannabinoid receptors in human immature oocytes and unfertilized metaphase-II oocytes». *Reproductive BioMedicine Online*. **23**(3), 372-379.

[12] López-Cardona, A. P., Sánchez-Calabuig, M. J., Beltran-Breña, P., Agirregoitia, N., Rizos, D., Agirregoitia, E. eta Gutierrez-Adán, A. 2016. «Exocannabinoids effect on in vitro bovine oocyte maturation via activation of AKT and ERK1/2». *Reproduction*. **152**(6), 603-612.

[13] López-Cardona, A. P., Pérez-Cerezales, S., Fernández-González, R., Laguna-Barraza, R., Pericuesta, E., Agirregoitia, N. eta Agirregoitia, E. 2017. «CB1 cannabinoid receptor drives oocyte maturation and embryo development via PI3K/Akt and MAPK pathways». *FASEB Journal*. **31**(8), 3372-3382.

[14] Agirregoitia, E., Ibarra-Lecue, I., Totorikaguena, L., Mendoza, R., Expósito, A., Matorras, R. eta Agirregoitia, N. 2015. «Dynamics of expression and localization of the cannabinoid system in granulose cells during oocyte nuclear maturation». *Fertility and Sterility*. **104**(3), 753-760.

[15] Agirregoitia, E., Totorikaguena, L., Expósito, A., Mendoza, R., Matorras, R. eta Agirregoitia, N. 2016. «Dynamic of expression and localization of cannabinoid-degrading enzymes FAAH and MGLL in relation to CB1 during meiotic maturation of human oocytes». *Cell and Tissue Research*. 1-9.

[16] Gammon, C. M., Freeman, G. M., Jr, Xie, W., Petersen, S. L., Wetsel, W. C., Petersen, S. L. eta Wetsel,W. C. 2005. «Regulation of gonadotropinreleasing hormone secretion by cannabinoids». *In Endocrinology*. **146**, 4491-4499.

[17] Smith, C. G., Besch, N. F., Smith, R. G., Besch, P.K. 1979. «Effect of tetrahydrocannabinol on hypothalamic -pituitary axis in ovariectomized rhesus monkey». *Fertility and Sterility*. **31**, 335-339.

[18] Walker, O., Holloway, A. eta Raha, S. 2019. «The role of the endocannabinoid system in female reproductive tissues». *Journal of Ovarian Research*. **12**(Suppl. 1), 2-10.

[19] El-Talatini, M. R., Taylor, A. H. eta Konje, J. C. 2009. «Fluctuation in anandamide levels from ovulation to early pregnancy in in-vitro fertilization-embryo transfer women, and its hormonal regulation». *Human Reproduction*. **24**(8), 1989-1998.

[20] Bambang, K. N., Karasu, T., Gebeh, A., Taylor, A. H., Marczylo, T. H., Lam, P., Willets, J. M. eta Konje, J. C. 2010. «From fertilization to implantation in mammalian pregnancy-modulation of early human reproduction by the endocannabinoid system». *Pharmaceuticals (Basel)*. **3**, 2910-2929.

[21] Cacciola, G., Chioccarelli, T., Altucci, L., Ledent, C., Mason, J. I., Fasano, S. eta Cobellis, G. 2013. «Low 17beta-estradiol levels in Cnr1 knock-out mice affect spermatid chromatin remodeling by interfering with chromatin reorganization». *Biology of Reproduction*. **88**(6), 152.

[22] Wenger, T., Ledent, C., Csernus, V. eta Gerendai, I. 2001. «The central cannabinoid receptor inactivation suppresses endocrine reproductive functions». *Biochemical and Biophysical Research Communications*. **284**(2), 363-368.

[23] Oláh, M., Milloh, H. eta Wenger, T. 2008. «The role of endocannabinoids in the regulation of luteinizing hormone and prolactin release. Differences between the effects of AEA and 2AG». *Molecular and Cellular Endocrinology*. **286**(1-2 Suppl. 1), 36-40.

[24] Albertini, D.F., Combelles, C.M., Benecchi, E eta Carabatsos, M.J. 2001. «Cellular basis for paracrine regulation of ovarian follicle development». *Reproduction*. **121**, 647-53.

[25] Asch, R.H., Smith, C.G., Siler-Khodr, T.M. eta Pauerstein, C.J. 1981. «Effects of delta 9-tetrahydrocannabinol during the follicular phase of the rhesus monkey (macaca mulatta)». *J Clin Endocrinol Metab*. **1**, 50-5.

[26] Wang, H., Guo, Y., Wang, D., Kingsley, P. J., Marnett, L. J., Das, S. K. eta Dey, S. K. 2004. «Aberrant cannabinoid signaling impairs oviductal transport of embryos». *Nature Medicine*. **10**(10), 1074-1080.

[27] Totorikaguena, L., Olabarrieta, E., Lolicato, F., Romero Aguirregomezcorta, J., Smitz, J., Agirregoitia, N. eta Agirregoitia, E. 2020. «The endocannabinoid system modulates the ovarian physiology and its activation can improve *in vitro* oocyte maturation». *Journal of Cellular Physiology*. doi: 10.1002/jcp.29663

[28] Shi, L., Zhang, J., Lai, Z., Tian, Y., Fang, L., Wu, M. eta Wang, S. 2016. «Longterm moderate oxidative stress decreased ovarian reproductive function by reducing follicle quality and progesterone production». *PLOS One*. **11**(9), 1-18.

[29] Otoi, T., Yamamoto, K., Koyama, N., Tachikawa, S eta Suzuki, T. 1997. «Bovine oocyte diameter in relation to developmental competence». *Theriogenology*. **48**(5), 769-74.

[30] Totorikaguena, L., Olabarrieta, E., López-Cardona, A. P., Agirregoitia, N. eta Agirregoitia, E. 2019. «Tetrahydrocannabinol modulates in vitro maturation of oocytes and improves the blastocyst rates after in vitro fertilization». *Cellular Physiology and Biochemistry*. 53(3), 439-452.

Egunero ariketa fisiko labur baina bizia egiteak nabarmen hobetzen du osasuna

Orain arte jakinekoa ez bazen ere, egunero zenbait aldiz ariketa fisiko labur-laburrak baina biziak egitea estrategia oso eraginkorra da osasuna hobetzeko. Minbiziagatik edo gaitz kardiobaskularrengatik hiltzeko arriskua % 40 baino gehiago murrizten du. Garrantzitsua da datu hori osasun-politiketan aintzat hartzea eta bizimodu aktiboa sustatzeko baliatzea.

Orain arte, Osasunaren Mundu Erakundeak astero 150-300 minutuz ariketa fisiko lasaia edo 75-150 minutuz ariketa fisiko bizia egitea gomendatu izan du, osasun orokorra zaintze aldera. Alabaina, azken ikerketek frogatu dute ariketa bizi oso laburrak ere lagungarri direna osasuntsu egoteko.

Bizimodu pasiboa duten 25.000 pertsonari baino gehiagori jarraipena eginez, ikusi dute egunero ariketa biziko minutu bateko edo biko hiru saio labur txertatze soilak % 38-40 murrizten duela minbiziagatik eta bestelako arrazoi batzuengatik hiltzeko arriskua, eta % 48-49 gaitz kardiobaskularren eraginez hiltzekoa. Eguneroko bizitzan erraz egin daitezkeen ariketak dira: eskailerak bizi igotzea, azkar ibiltzea eta pisuak garraiatzea, besteak beste. Ariketa fisikoa gimnasioan, kirol batean zein bestelako entrenamenduetan modu egituratu batean egitea bezain eraginkorra omen da, ikerketaren arabera.

Ariketa fisiko biziak inflamazioa murriztu eta bihotz-biriketako gaitasuna hobetzen du, eta horrek ondorio zuzenak ditu organismoan: bihotz-biriketako gaitasuna unitate metaboliko baten baliokidea hobetzen den bakoitzeko, % 7 murrizten da gaitz kardiobaskularrak garatzeko eta minbiziagatik hiltzeko arriskua. Bularreko, endometrioko eta koloneko minbizietan, besteak beste.

Kirol-erlojuei esker jakin da hori guztia. Orain arte, 10 minututik gorako ariketa fisikoa bakarrik hartu da aintzat, eta aisialdian egiten diren kirol-ekintzekin lotu izan da. Eskumuturreko azelerometroen teknologia eramangarria garatzeak, ordea, agerian jarri du zenbateko eragina duten egunerokotasunean egiten diren ariketa laburrek.

Bizitza sedentarioari aurre

Emaitzek inplikazio handiak dituzte osasun publikoaren alorrean. 40 urtetik gorako heldu gehienek bizitza sedentarioa izaten dute, askori ez baitzaie erakargarri egiten ariketa fisiko egituratua, edo ezin baitute halakorik egin. Ondorioz, haien osasunak okerrera egiten du zahartu aurretik.

Egunerokotasunean ariketa bizi intermitenteak txertatzea, ordea, eskuragarria da edozeinentzat, ez baitu eskatzen prestakuntza fisikorik, ez instalazio berezirik ezta denbora-inbertsiorik ere. Nahikoa da egunerokotasunak eskatzen dituen ariketak modu kontzientean egitea.

Ikertzaileen ustez, garrantzitsua da osasun publikoko politikek eta gizarte-politikek aintzat hartzea emaitza horiek. Herritarrek, erosotasunaren izenean, neke-momentu txikiak dakartzaten jarduerak ekiditeko joera izaten dute egunerokoan, eta garrantzitsua da jarduera bizi horiek berriz integratzeko jarraibideak ematea. Esaterako, igogailua baztertu eta eskailerak igotzea, erosketetako poltsak garraiatzea edo etxeko lan biziak berreskuratzea. Intentsitate handiko jarduera horiek osasuna ematen dutela ulertzea biztanleriaren osasun orokorra hobetu eta gaitz larriak ekiditeko estrategia moduan baliatu daiteke.

«Egunerokotasunean minutu bateko ariketa biziak txertatzeak % 38-40 murrizten du minbiziagatik hiltzeko arriskua»

InpAkta 00 (2023) 107-124

https://doi.org/10.1038/s41591-022-02100-x

Association of wearable device-measured vigorous intermittent lifestyle physical activity with mortality

Emmanuel Stamatakis[1*], Matthew N. Ahmadi[1], Jason M. R. Gill[2], Cecilie Thøgersen-Ntoumani[3], Martin J. Gibala[4], Aiden Doherty[5] & Mark Hamer[6]

[1] Charles Perkins Centre, Faculty of Medicine and Health, The University of Sydney; [2] School of Cardiovascular and Metabolic Health, University of Glasgow; [3] Danish Centre for Motivation and Behaviour Science, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark; [4] Department of Kinesiology, McMaster University; [5] Big Data Institute, Nuffield Department of Population Health, University of Oxford; [6] Institute Sport Exercise Health, Division Surgery Interventional Science, University College London

ABSTRACT: Wearable devices can capture unexplored movement patterns such as brief bursts of vigorous intermittent lifestyle physical activity (VILPA) that is embedded into everyday life, rather than being done as leisure time exercise. Here, we examined the association of VILPA with all-cause, cardiovascular disease (CVD) and cancer mortality in 25,241 nonexercisers (mean age 61.8 years, 14,178 women/11,063 men) in the UK Biobank. Over an average follow-up of 6.9 years, during which 852 deaths occurred, VILPA was inversely associated with all three of these outcomes in a near-linear fashion. Compared with participants who engaged in no VILPA, participants who engaged in VILPA at the sample median VILPA frequency of 3 length-standardized bouts per day (lasting 1 or 2 min each) showed a 38%–40% reduction in all-cause and cancer mortality risk and a 48%–49% reduction in CVD mortality risk. Moreover, the sample median VILPA duration of 4.4 min per day was associated with a 26%–30% reduction in all-cause and cancer mortality risk and a 32%–34% reduction in CVD mortality risk. We obtained similar results when repeating the above analyses for vigorous physical activity (VPA) in 62,344 UK Biobank participants who exercised (1,552 deaths, 35,290 women/27,054 men). These results indicate that small amounts of vigorous nonexercise physical activity are associated with substantially lower mortality. VILPA in nonexercisers appears to elicit similar effects to VPA in exercisers, suggesting that VILPA may be a suitable physical activity target, especially in people not able or willing to exercise.

* **Corresponding author / Harremanetan jartzeko:** Emmanuel Stamatakis, Charles Perkins Centre, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia. e-mail: emmanuel.stamatakis@sydney.edu.au. **Contributed by** Jon Irazusta **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Stamatakis, Emmanuel *et al.* «Association of wearable device-measured vigorous intermittent lifestyle physical activity with mortality», *Nature Medicine* | Volumen 28 | December 2022:2521–2529 (https://doi.org/10.1038/s41591-022-02100-x)

Physical activity is associated with reduced mortality risk[1], and reduced risk of CVD[1] and certain cancers [2-4]. Recently updated guidelines [4, 5], based mostly on questionnaire-derived evidence, recommend 150-300 min of moderate-intensity activity or 75-150 min of vigorous-intensity physical activity (≥6 metabolic equivalents) per week. New emphasis is placed on "all activity counts" occurring across all life domains and regardless of bout duration. This recommendation contrasts with previous guide-lines [6, 7] that did not recognize the health value of physical activity bouts lasting <10 min. Besides, little evidence supports the previous guideline because questionnaires can typically capture only longer bouts (for example, ≥10 min) of physical activity and often concentrate on leisure time activities such as gym-based exercise, running and sports [8, 9].

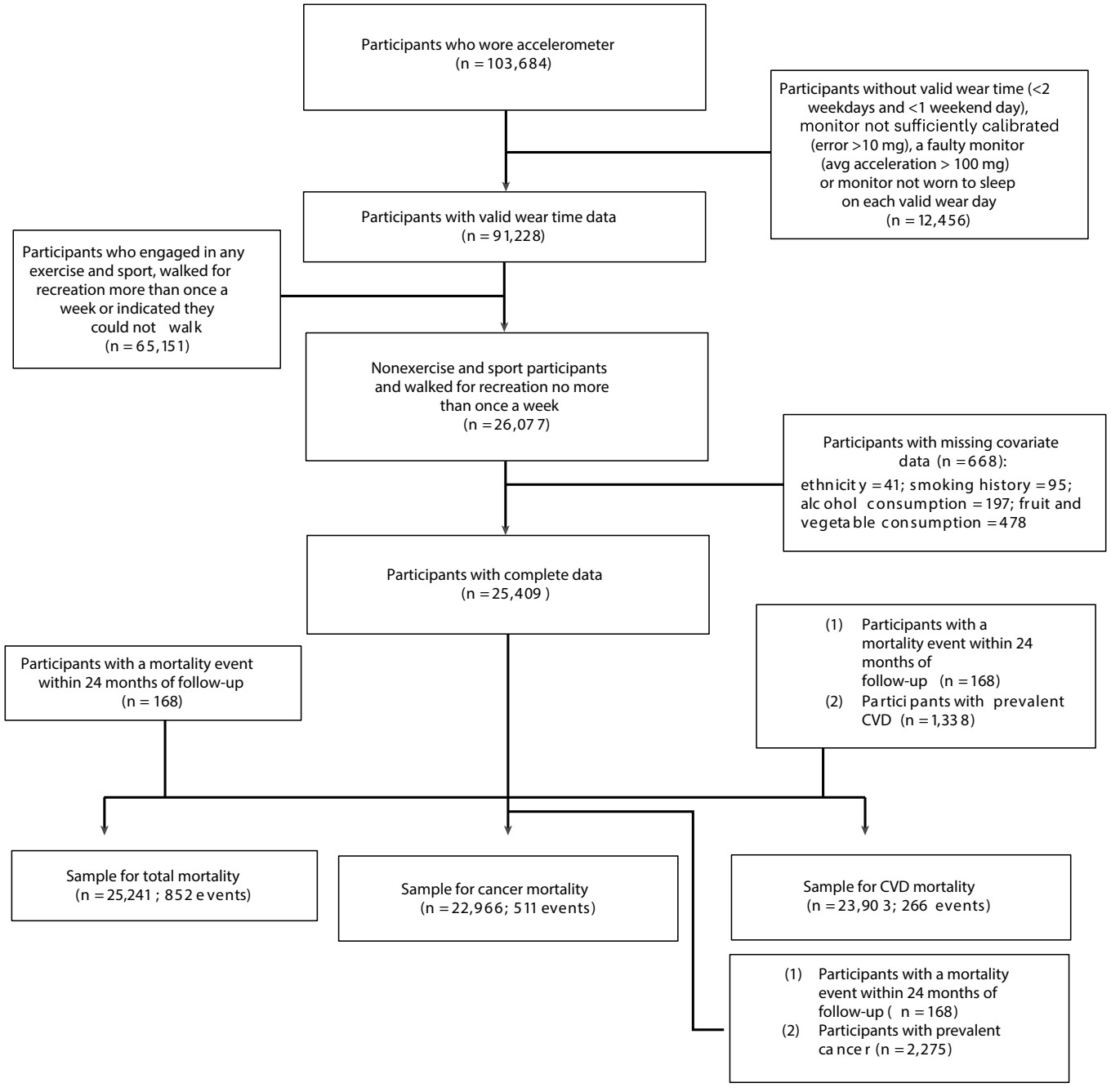


Figure 1. Flow diagram of non-exercisers. Flow diagram of UK Biobank participants for the dose–response analyses of VILPA

The health effects of each time unit of physical activity are intensity dependent [10-13]. For a given volume of physical activity, higher contributions of VPA are associated with additional mortality risk reduction [10-13]. This is partly due to the enhanced cardiorespiratory adaptations it causes [14] and the protection it offers against the development of certain cancers [15, 16]. Although vigorous-intensity physical activity is time-efficient, vigorous structured exercise-based sessions (for example, gym-based, sports, high-intensity interval training) are not feasible or appealing to the majority of middle-aged adults, as indicated by the very low participation rates [8, 9, 17]. Over a median follow-up of 3.1 years, a previous UK Biobank accelerometry study [18] concluded that moderate-to vigorous-intensity activity was associated with greater reductions in all-cause mortality risk than lower-intensity activity. However, VPA was not specifically quantified in this study [18].

Table 1
Sample characteristics, as stratified by average daily frequency of VILPA bouts (*n* = 25,241)

	Number of daily VILPA bouts lasting up to 2minutes				
	0	1-2	3-4	≥4	Overall
n	2,816	8,088	8,768	5,569	25,241
Follow-up, years	6.8 (1.0)	6.9 (0.8)	6.9 (0.8)	6.9 (0.7)	6.9 (0.8)
Age, mean (s.d.)	64.7 (6.8)	62.9 (7.4)	61.3 (7.6)	59.7 (7.7)	61.8 (7.6)
Male, *n* (%)	888 (31.5)	2,977 (36.8)	4,079 (46.5)	3,119 (56.0)	11,063 (43.8)
Ethnicity, *n* (%)					
Asian	30 (1.1)	93 (1.1)	119 (1.4)	84 (1.5)	326 (1.3)
Black	18 (0.6)	74 (0.9)	89 (1.0)	80 (1.4)	261 (1.0)
Mixed	16 (0.6)	40 (0.5)	55 (0.6)	46 (0.8)	157 (0.6)
Other	21 (0.7)	76 (0.9)	66 (0.8)	60 (1.1)	223 (0.9)
White	2,731 (97.0)	7,805 (96.5)	8,439 (96.2)	5,299 (95.2)	24,274 (96.2)
Smoking history, *n* (%)					
Current	328 (11.6)	766 (9.5)	775 (8.8)	456 (8.2)	2,325 (9.2)
Previous	1,042 (37.0)	2,898 (35.8)	3,100 (35.4)	1,905 (34.2)	8,945 (35.4)
Never	1,446 (51.3)	4,424 (54.7)	4,893 (55.8)	3,208 (57.6)	13,971 (55.4)
Body mass index	29.6 (5.9)	28.4 (5.3)	27.3 (4.7)	26.3 (4.3)	27.6 (5.1)
Alcohol consumption, *n* (%)[a]					
Never	138 (4.9)	345 (4.3)	300 (3.4)	190 (3.4)	973 (3.9)
Ex-drinker	140 (5.0)	307 (3.8)	266 (3.0)	155 (2.8)	868 (3.4)
Within guidelines	1,760 (62.5)	4,939 (61.1)	5,087 (58.0)	3,203 (57.5)	14,989 (59.4)
Above guidelines	778 (27.6)	2,497 (30.9)	3,115 (35.5)	2,021 (36.3)	8,411 (33.3)
Education, *n* (%)					
College	975 (35.3)	2,941 (37.0)	3,196 (37.2)	2,002 (36.7)	9,114 (36.8)
A/AS level	380 (13.8)	1,010 (12.7)	1,107 (12.9)	702 (12.9)	3,199 (12.9)
O level	621 (22.5)	1,763 (22.2)	1,955 (22.7)	1,209 (22.2)	5,548 (22.4)
CSE	105 (3.8)	376 (4.7)	444 (5.2)	360 (6.6)	1,285 (5.2)
NVQ/HND/HNC	147 (5.3)	461 (5.8)	557 (6.5)	401 (7.4)	1,566 (6.3)
Other	535 (19.4)	1,389 (17.5)	1,337 (15.6)	776 (14.2)	4,037 (16.3)

	Number of daily VILPA bouts lasting up to 2minutes				
	0	1-2	3-4	≥4	Overall
Fruit and vegetable consumption, *n* (%)[b]					
High	889 (31.6)	2,517 (31.1)	2,693 (30.7)	1,689 (30.3)	7,788 (30.9)
Moderate	1,287 (45.7)	3,766 (46.6)	4,067 (46.4)	2,593 (46.6)	11,713 (46.4)
Low	640 (22.7)	1,805 (22.3)	2,008 (22.9)	1,287 (23.1)	5,740 (22.7)
Family history of CVD, *n* (%)	1,680 (59.7)	4,588 (56.7)	4,752 (54.2)	2,919 (52.4)	13,939 (55.2)
Family history of cancer, n (%)	748 (26.6)	2,109 (26.1)	2,221 (25.3)	1,366 (24.5)	6,444 (25.5)
Medication, *n* (%)					
Cholesterol	720 (25.6)	1,613 (19.9)	1,339 (15.3)	678 (12.2)	4,350 (17.2)
Blood pressure	891 (31.6)	1,859 (23.0)	1,639 (18.7)	740 (13.3)	5,129 (20.3)
Insulin	54 (1.9)	87 (1.1)	65 (0.7)	28 (0.5)	234 (0.9)
Self-rated health, *n* (%)					
Poor	240 (8.5)	491 (6.1)	350 (4.0)	142 (2.5)	1,223 (4.8)
Fair	868 (30.8)	1,969 (24.3)	1,845 (21.0)	1,030 (18.5)	5,712 (22.6)
Good	1,433 (50.9)	4,630 (57.2)	5,290 (60.3)	3,432 (61.6)	14,785 (58.6)
Excellent	259 (9.2)	965 (11.9)	1,270 (14.5)	956 (17.2)	3,450 (13.7)
Sleep (hours per day), median [IQR]	7.3 [6.3, 8.2]	7.3 [6.3, 8.1]	7.4 [6.5, 8.1]	7.4 [6.5, 8.1]	7.4 [6.4, 8.1]
Acceleration magnitude (milli-gravity) [IQR]	20.7 [17.2, 25.3]	24.5 [20.9, 29.3]	28.2 [24.3, 33.3]	33.5 [28.8, 39.6]	27.2 [22.4, 30.6]
Total activity (min per day), median [IQR][c]	110.2 [70.5, 169.5]	119.1 [82.7, 181.3]	136.3 [99.9, 196.8]	176.8 [133.4, 239.9]	138.3 [96.0, 201.0]
Light activity (min per day), median [IQR]	92.8 [59.0, 145.4]	94.6 [63.8, 146.7]	98.0 [69.0, 146.6]	105.2 [75.4, 149.2]	98.0 [68.0, 147.0]
Moderate activity (min per day), median [IQR]	12.8 [6.4, 24.3]	20.0 [11.4, 33.9]	27.9 [17.3, 44.1]	39.7 [26.0, 60.0]	25.9 [14.6, 43.2]
Vigorous activity (min per day), median [IQR]	—	1.6 [0.9, 2.3]	4.7 [2.6, 5.9]	8.1 [7.3, 9.3]	4.0 [1.3, 9.1]
Percent of total activity in vigorous activity [IQR]	—	0.8 [0.3, 1.7]	3.7 [2.0, 6.1]	7.3 [3.9, 13.8]	3.2 [1.0, 7.9]
VILPA bouts frequency (up to 1 min duration), median [IQR]	—	1 [1, 2]	3 [3, 4]	7 [6, 9]	3 [2, 4]
VILPA bouts frequency (up to 2 min duration), median [IQR]	—	1 [1, 2]	3 [3, 4]	8 [6, 10]	3 [2, 4]
Mortality rate (per 1,000 person-years)					
All-cause mortality	10.4	5.2	4.2	2.6	4.9
CVD mortality[d]	3.1	1.8	1.3	0.5	1.5
Cancer mortality[e]	7.3	3.4	2.8	1.6	3.2

The columns breakdown corresponds to length-standardized VILPA bouts. Values represent mean (s.d.) unless specified otherwise. A/AS level, ; CSE, ; Higher National Certificate, ; Higher National Diploma, ; IQR, interquartile range; National Vocational Qualification, ; O level, . [1]Alcohol consumption: above guidelines is >14 units per week, where 1 unit = 8 g of ethanol. [2]Fruits and vegetable consumption: low is <5 servings per day, high is >8 servings per day. [3]Daily duration of light-, moderate- and vigorous-intensity activity. [4]Calculated from the CVD mortality sample (n = 23,903). [5]Calculated from the cancer mortality sample (n = 22,699).

VILPA [19] refers to brief and sporadic (for example, up to 1 or up to 2 min long) bouts of vigorous-intensity physical activity done as part of daily living, such as bursts of very fast walking while commuting to work or moving from place to place, or stair climbing [20]. No cohort study has examined the associations of VILPA with mortality or other prospective outcomes. For most adults, VILPA may be more feasible than structured exercise because it requires minimal time commitment and involves no specific preparation, equipment or access to facilities. Many common activities of daily living are likely to elicit relative vigorous-intensity effort in physically inactive adults with poor fitness who do not habitually exercise [21], which is the majority demographic in many countries [8, 9, 22, 23].

In contrast to questionnaires, wearable devices such as wrist [24, 25] or thigh [26, 27] accelerometers continuously record movement at a high resolution allowing them to capture fine-grain patterns of brief physical activity bouts, such as VILPA [19]. The rapidly growing use of wearable devices in research [25, 27-30] and among consumers [31] offers opportunities to better understand the health-enhancing potential of VILPA and analogous unexplored movement "micro-patterns". Such potential is greatly enhanced by the recent application of machine learning [32-34] in studies using wearable devices to understand the health effects of movement. In a sample of UK Biobank participants with accelerometry data who reported no exercise in their leisure time, we examined the dose– response curves and minimum VILPA dose (daily duration and bout frequency) associated with all-cause, CVD and cancer mortality risk. To provide a population-wide context for our findings, we also examined the dose–response associations of (exercise or nonexercise) VPA with the same mortality outcomes among exercisers in the UK Biobank accelerometry substudy.

1. Results

1.1. Description of the study sample

Figure 1 shows the sample derivation process, which resulted in 25,241 (all-cause mortality analyses)/23,903 (CVD mortality analyses)/22,699 (cancer mortality analyses) UK Biobank participants being included in the corresponding analyses. Table 1 presents the characteristics of the sample by daily VILPA frequency. The mean (s.d.) age of participants was 61.8 (7.6) years, and 56.2% were female. Over a mean follow-up of 6.9 (0.8) years (175,528 person-years), 852 deaths were recorded (266 due to CVD and 511 due to cancer).

Supplementary Fig. 1 describes the sample derivation process for the exercisers sample (defined as those who reported any leisure time exercise/sports or more than one recreational walk per week). Over a mean follow-up of 6.9 (0.8) years (432,545 person-years) 62,344 exercisers were included in the all-cause mortality analyses (1,552 events), 56,810 were included in the CVD mortality analyses (303 events) and 56,397 were included in the cancer mortality analyses (736 events). Supplementary Table 1 describes the characteristics of exercisers who, in comparison with the nonexercisers, had higher educational attainment (46.3% versus 36.8% with college/university degree), higher self-rated health (25.2% versus 13.7% with excellent health) and lower medication use (for example, 15.3% versus 20.3% taking blood pressure medication).

1.2. VILPA summary and nonexerciser status

To enable examination of VILPA in our study (brief bouts of nonexercise VPA occurring during daily living), we used information on exercise participation available in the UK Biobank Study (Supplementary Table 2). Our core VILPA analyses only included 25,241 participants who at the UK Biobank baseline (on average 5.5 years before the accelerometry baseline) reported no leisure time exercise participation and no more than one recreational walk per week. For use in sensitivity analyses, we also derived an alternative, more conservative, definition of nonexercisers by excluding participants who reported any recreational walking in addition to any leisure time exercise (n = 10,230). A subsample analysis among 2,407 participants of our core sample who had a UK Biobank re-examination an average (s.d.) of 1.5 (1.4) years before the accelerometry measurements showed that the nonexerciser status was stable over time: 82% reported no leisure time physical activity and no more than one recreational walking session per week on both time points. Among the 6,095 entire UK Biobank accelerometry sample participants who reported no exercise at baseline and had a re-examination, 88% maintained their no leisure time physical activity status over time.

In the core VILPA analyses sample of 25,241 participants, almost all VILPA was accrued in bouts lasting up to 1 or up to 2 min: 92.3% of bouts lasted up to 1 min and 97.7% lasted up to 2 min. Excluding VILPA values of zero, the median and maximum VILPA daily duration was 4.0 and 16.0 min per day for both bout lengths; the median and maximum VILPA frequency was 3.0 and 11.0 length-standardized bouts per day.

Among the 62,344 exercisers entered in the comparative analyses, the large majority of context-agnostic (exercise or nonexercise) VPA was accrued in bouts lasting up to 2 min (93.1% of all VPA bouts). Median and maximum VPA daily duration was 6.2 and 18.0 min per day; the median and maximum daily frequency was 4.4 and 14.0 length-standardized bouts per day. In the nonexercisers sample, 11.2% of participants recorded no VILPA. In the exercisers sample, 6.9% recorded no VPA.

1.3. Associations of VILPA with all-cause mortality

In multivariable-adjusted analyses (adjusted for age, sex, light- and moderate-intensity physical activity, longer VPA bouts, smoking, alcohol, sleep duration [35, 36], fruit and vegetable consumption, education, parental history of CVD and cancer, medication use, and prevalent CVD and cancer; Supplementary Table 3), bouts lasting up to 1 min (Fig. 2a,c) and up to 2 min (Fig. 2b,d), exhibited a near —linear dose— response associations of daily VILPA daily duration and frequency with all-cause mortality. Supplementary Table 4A presents the hazard ratio (HR) and 95% confidence intervals (CI) associated with the minimum dose (eliciting 50% of the total effect) [37, 38], and the median and maximum VILPA daily duration and frequency for each bout length. The minimum frequency dose for length-standardized VILPA bouts lasting 1 min was 1.5 bouts per day corresponding to a HR of 0.75 (95% CI 0.66, 0.85). The median and maximum VILPA frequency for length-standardized bouts lasting 1 min were associated with a HR of 0.61 (0.50, 0.74) and 0.52 (0.37, 0.72), respectively. The minimal daily duration dose [37, 38] for VILPA bouts lasting up to 1 min was 3.4 min per day corresponding to a HR of 0.78 (95% CI 0.70, 0.86). The median and maximum VILPA volumes for bouts lasting up to 1 min were associated with a HR of 0.73 (0.63,0.85) and 0.59 (0.46, 0.74). All-cause mortality findings for bouts lasting up to 2 min were similar in terms of the dose–response curves (Fig. 2b,d), the minimal dose values and the magnitude of the associations linked to the median and maximum VILPA daily duration and frequency (Supplementary Table 4A).

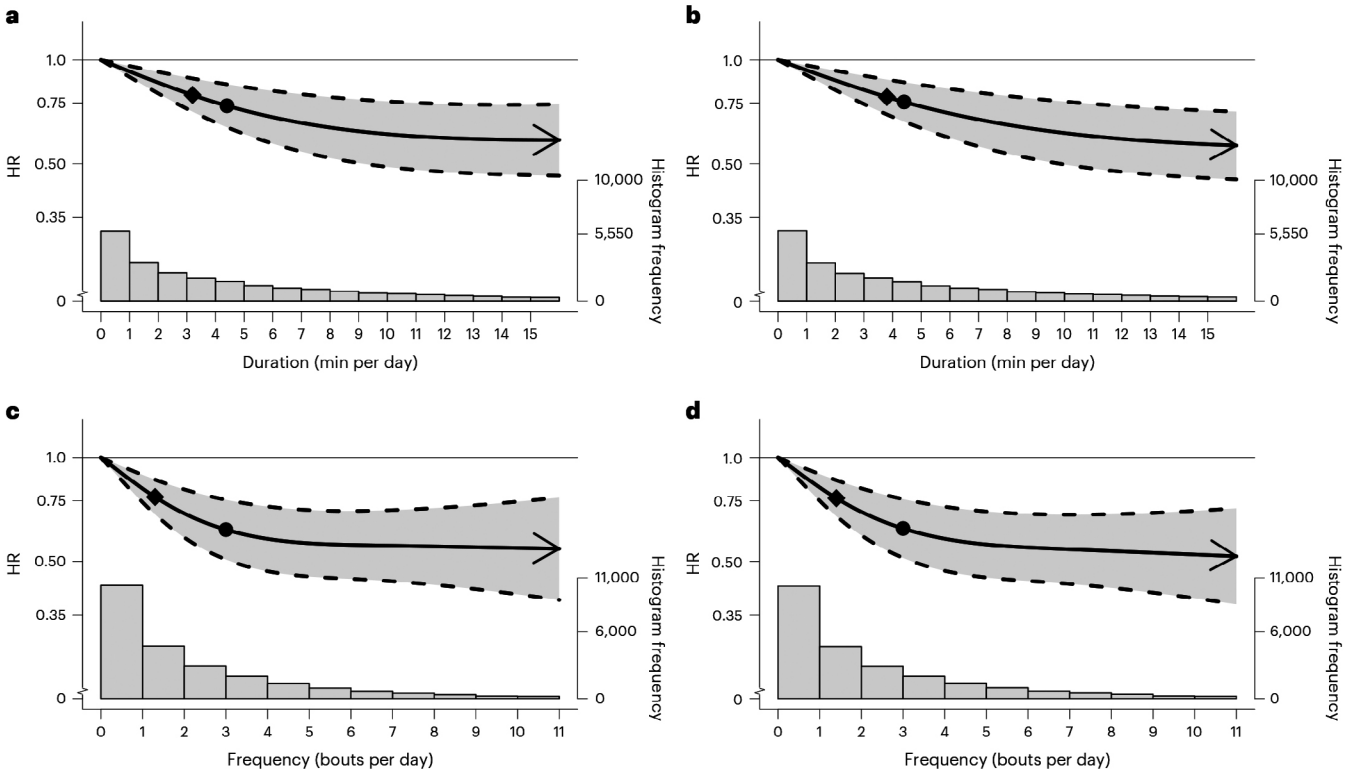


Figure 2. Association of the daily duration and frequency of VILPA with all-cause mortality. a,b, Dose–response curves showing all-cause mortality

HR associated with increasing daily duration of VILPA, for bouts of VILPA up to 1 min (a) and 2 min (b) in duration. c,d, Dose–response curves showing all-cause mortality HR associated with increasing daily frequency of VILPA, for length-standardized bouts of VILPA 1 min (c) and 2 min (d) in duration. Data are shown for *n* = 25,241 participants with 852 events and with a mean followup of 6.9 (0.8) years. Diamond, minimal dose, as indicated by the ED50 statistic which estimates the daily duration/frequency of VILPA associated with 50% of optimal risk reduction. Circle, HR associated with the median VILPA value (see Supplementary Table 4 for the list of values). Data are adjusted for the covariates listed in the online Methods. The shaded region demarcated by dashed lines represents the 95% CI. The solid line that lies within the shaded region represents the HR. The arrowhead represents the absence of an observed inflection point (for example, larger risk reduction with higher amounts of VILPA). The histogram on the right shows the sample distribution.

1.4. Associations of VILPA with CVD mortality

The beneficial associations found in the CVD mortality multivariable-adjusted analyses were more pronounced than the all-cause mortality findings for both bout lengths (Fig. 3a-d and Supplementary Table 4B). For example, the minimum frequency dose for length-standardized VILPA bouts lasting 1 min was 1.4 bouts per day corresponding to a HR of 0.67 (95% CI 0.52, 0.86), and the median and maximum VILPA frequency were associated with a HR of 0.51 (0.35, 0.74) and 0.35 (0.15, 0.81), respectively. The minimal CVD mortality daily duration dose for VILPA bouts lasting up to 1 min was 3.4 min per day corresponding to a HR of 0.73 (95% CI 0.58, 0.91). The median and maximum VILPA daily duration values were associated with a HR of 0.66 (0.50, 0.88) and 0.45 (0.29, 0.72), respectively. CVD mortality findings for bouts lasting up to 2 min were very similar in terms of the dose–response curves (Fig. 3b,d) and all other metrics (Supplementary Table 4B).

1.5. Associations of VILPA with cancer mortality

The findings of the cancer mortality multivariable-adjusted analyses were very consistent with the equivalent all-cause mortality analyses outlined above, in terms of both the dose–response curves (Fig. 4a-d) and the point estimates associated with the minimum dose and the median and maximum VILPA frequency and daily duration values (Supplementary Table 4C). For example, the minimum frequency dose for length-standardized VILPA bouts lasting 1 min was 1.5 bouts per day corresponding to a HR of 0.75 (95% CI 0.63, 0.88). The minimal cancer mortality daily duration dose for VILPA bouts lasting up to 1 min was 3.4 min per day corresponding to a HR of 0.76 (95% CI 0.66, 0.87), whereas the median and maximum VILPA daily duration values were associated with a HR of 0.70 (0.59, 0.84) and 0.51 (0.38, 0.69). Like the other two mortality outcomes, cancer mortality findings for bouts lasting up to 2 min were very similar to bouts lasting up to 1 min across all metrics (Fig. 4a-d and Supplementary Table 4C).

1.6. Sensitivity analyses

Excluding participants with poor health (*n* = 1,223) and additionally adjusting for body mass index (Extended Data Figs. 1-3) did not appreciably change the results.

E-values indicated that for our estimates to be null the association of an unmeasured confounder with exposures and mortality should be a HR (lower 95% CI) of 1.87 (1.54) to 3.26 (2.12) for all-cause mortality; 2.10 (1.44) to 5.16 (1.77) for CVD mortality; or 1.97 (1.56) to 3.50 (2.00) for cancer mortality (Supplementary Table 5).

Categorical analyses of VILPA daily duration (Extended Data Fig. 4a,b) and frequency (Extended Data Fig. 4c,d) by VILPA tertile-based groups produced results consistent with the main dose–response analyses. Similarly, restricting analyses to those who reported no recreational walking and no leisure time exercise (*n* = 10,230) produced results that were very consistent with the main results in the core (*n* = 25,241) VILPA sample (Supplementary Fig. 2).

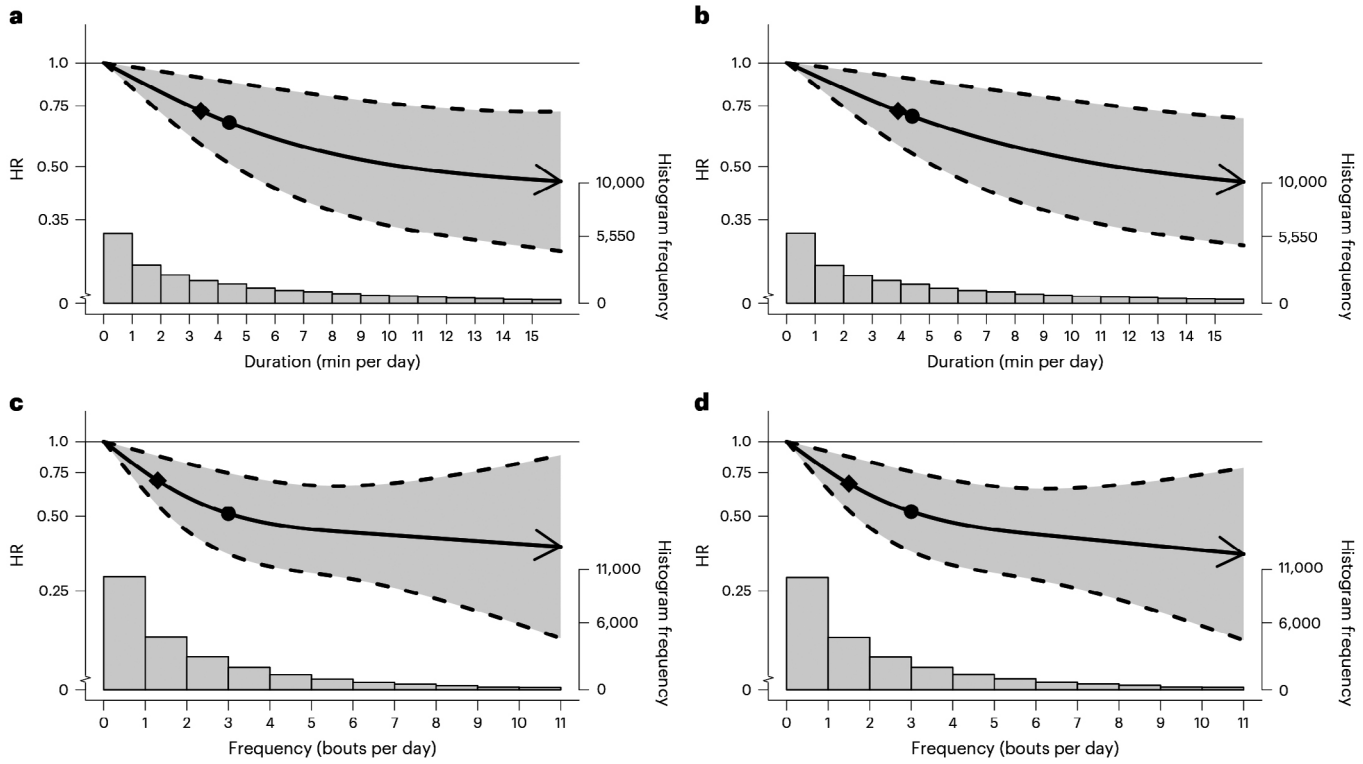


Figure 3. Association of the daily duration and frequency of VILPA with CVD mortality. a,b, Dose-response curves showing CVD mortality HRs associated with increasing daily duration of VILPA, for bouts of VILPA up to 1 min (a) and 2 min (b) in duration. c,d, Dose-response curves showing CVD mortality HRs associated with increasing daily frequency of VILPA, for length-standardized bouts of VILPA1 min (c) and 2 min (d) in duration. Data are shown for ***n*** = 23,903 participants with 266 events and with a mean follow-up of 6.9 (0.8) years. Diamond, minimal dose, as indicated by the ED_{50} statistic which estimates the daily duration/frequency of VILPA associated with 50% of optimal risk reduction. Circle, HR associated with the median VILPA value (see Supplementary Table 4 for the list of values). Data are adjusted for the covariates listed in the online Methods. The shaded region demarcated by dashed lines represents the 95% CI. The solid line that lies within the shaded region represents the HR. The arrowhead represents the absence of an observed inflection point (for example, larger risk reduction with higher amounts of VILPA). The histogram on the right shows the sample distribution

1.7. Comparisons between nonexercisers (VILPA) and exercisers (VPA)

Context-agnostic (that is, exercise or nonexercise) VPA in exercisers exhibited an almost identical daily duration and frequency dose-response to VILPA in nonexercisers for all-cause mortality (Extended Data Figs. 5 and 6), with relatively modest differences in minimum dose (4.8 versus 3.4 min per day). No material differences in the CVD and cancer mortality dose-response curves were evident between the two strata and the corresponding minimum doses (Extended Data Figs. 5B,C and 6B,C). Similarly, no appreciable differences existed between exercisers' and nonexercisers' VPA/VILPA frequency in terms of their dose–response with the three mortality outcomes (Extended Data Fig. 6). Across all the above analyses there was an almost complete overlap of the 95% CI of the dose-response curves of the two strata/exposures. With very few exceptions, the point estimates associated with the minimum dose and the median VPA frequency and daily duration values (Supplementary Table 6) were also very similar to the equivalent VILPA data (Supplementary Table 4).

Volume analyses based on VILPA (nonexercisers) or VPA (exercisers) energy expenditure (kJ per kg per day) produced evidence of L-shaped associations with all three mortality outcomes, with steeper risk reductions in the lower end of the VILPA/VPA continuum (Supplementary Fig. 3). The only notable exception to this pattern was the

VILPA-CVD mortality curve among nonexercisers which indicated a linear association (Supplementary Fig. 3B). Data sparsity and a low number of events at higher levels of the VILPA energy expenditure makes between-strata comparisons and interpretation of these data challenging and less conclusive.

2. Discussion

Despite the large health potential of vigorous-intensity physical activity, most adults aged 40 and over do not do vigorous exercise or sports [8, 9, 17]. Our study is the first investigation, to the best of our knowledge, into the long-term health effects of nonexercise VPA embedded into daily living. We found consistent evidence of beneficial associations of relatively modest VILPA amounts with all-cause, CVD and cancer mortality. VILPA in nonexercisers appeared to elicit beneficial dose-response associations with mortality of similar magnitude to VPA in exercisers, a finding that emphasizes the potential of promoting higher intensity physical activity outside the leisure time exercise domain. Our work has relevance for the development of public health and clinical guidelines because people reporting no structured exercise in leisure time, yet recording VILPA bouts, may be unaware that they are taking short bouts of health-enhancing physical activity of higher intensity. Future guidelines could place emphasis on making people aware that they could potentially experience important benefits from VILPA even though they do not consider themselves to be formal "exercisers".

Although steeper mortality risk reductions occurred at the lower end of the VILPA distribution (up to roughly the median frequency and daily duration values, that is around 3-4 length-standardized bouts per day or 4-5 min per day), there were continuing mortal ity gains with more VILPA in a near-linear fashion across all three outcomes. With little variation between bouts lasting up to 1 or 2 min and across the three mortality outcomes, a minimum of 3.4-4.1 min of VILPA per day on average was associated with a 22%-28% reduction in mortality risk (compared with not doing VILPA). In terms of minimum daily frequency dose, fewer than two VILPA bouts (lasting 1 or 2 min) were associated with 24%-26% reduction in all-cause and cancer, and a 33% reduction in CVD mortality risk. The median VILPA frequency of 3 length-standardized bouts per day was associated with a 38%-40% reduction in all-cause and cancer mortality and a 48%-49% reduction in CVD mortality risk. The median daily VILPA duration of 4.4 min per day was associated with a 26%-30% reduction in all-cause and cancer mortality and a 32%-34% reduction in CVD mortality risk.

These results are striking but plausible. Proof-of-concept trials [39] have shown that very small doses of exercise-based intermittent VPA can have rapid and measurable effects on cardiorespiratory fitness, a key causal determinant of CVD [14]. High-intensity interval training [40] and studies of intermittent stair climbing [39] have shown that VPA bursts lasting 20 s to a few minutes, performed three to five times a day, can result in substantial improvements in cardiorespiratory fitness in previously inactive adults within a few weeks, providing a plausible physiological basis [14] for the associations we observed. Our comparisons with exercisers suggest that vigorous exertion is equally important and potentially beneficial for people who choose to be active during leisure time and those less able or willing to do so. Maintained or improved cardiorespiratory fitness owing to vigorous exertion (including VILPA) may partly explain the associations with cancer mortality that we observed: previous observational studies have estimated that a one metabolic equivalent unit higher cardiorespiratory fitness (3.5 ml of oxygen uptake per kg per min) is associated with a 7% reduction in total cancer mortality risk [41]. VPA has also been shown to specifically reduce risk of common cancer sites such as breast [15], endometrial [16] and colon [16].

The 32%-34% lower CVD mortality risk associated with the median VILPA duration of 4.4 min per day

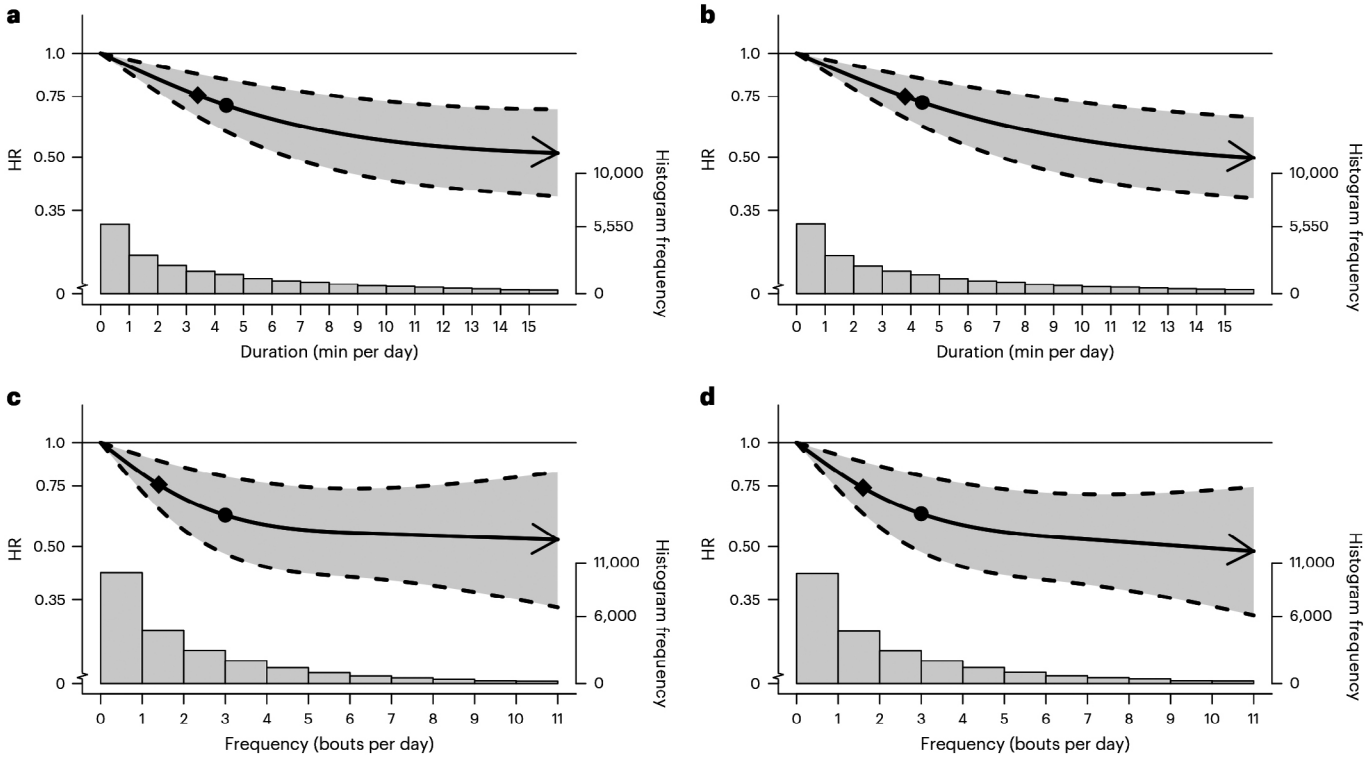


Figure 4. Association of the daily duration and frequency of VILPA with cancer mortality. a,b, Dose–response curves showing cancer mortality HRs associated with increasing daily duration of VILPA, for bouts of VILPA up to 1 min (a) and 2 min (b) in duration. c,d, Dose–response curves showing cancer mortality HRs associated with increasing daily frequency of VILPA, for length-standardized bouts of VILPA 1 min (c) and 2 min (d) in duration. Data are shown for ***n*** = 22,966 participants with 511 events and with a mean follow-up of 6.9 (0.8) years. Diamond, minimal dose, as indicated by the ED50 statistic which estimates the daily duration/frequency of VILPA associated with 50% of optimal risk reduction. Circle, HR associated with the median VILPA value (see Supplementary Table 4 for the list of values). Data are adjusted for the covariates listed in the online Methods. The shaded region demarcated by dashed lines represents the 95% CI. The solid line that lies within the shaded region represents the HR. The arrowhead represents the absence of an observed inflection point (for example, larger risk reduction with higher amounts of VILPA). The histogram on the right shows the sample distribution

(equivalent to just under 31 min of vigorous-intensity physical activity per week) that we observed is comparable with equivalent risk reduction for >75-150 min per week of questionnaire-measured vigorous leisure time physical activity reported in a recent US cohort (36%-45% lower risk compared with no leisure time vigorous activity done in bouts lasting at least 10 min) [11]. This seemingly sizeable difference in VPA amounts associated with a comparable effect size may be explained by the different measurements and domains employed in each study, the select sample of non-exercisers we employed in our study, and the strong possibility that the referent no vigorous leisure time physical activity group in this US study [11] may do some VILPA. Questionnaires [10-12] can only capture con tinuous blocks of time containing a mixture of vigorous activity with interruptions and rest, rather than actual time in vigorous intensity that the wearable devices in our study could quantify. Only one or two in five UK middle-aged adults engage in structured vigorous exercise at least once a month [8, 9, 17], suggesting numerous participation barriers. Our findings highlight the potential value of short VPA bursts during daily living to improve overall and cardiovascular health and reduce risk of cancer.

This is the first study of VILPA and prospective health outcomes, using device-based measurement and machine learning-based methods. Although we cannot entirely rule out reverse causation bias, our results were very robust to relevant

sensitivity analyses. E-values indicated that unmeasured confounding is unlikely to explain the associations we observed. Although some VILPA activities (for example, carrying heavy shopping bags) may not be perfectly captured by wrist-worn accelerometers, such measurement error is likely random leading to underestimation of the "true" associations with mortality, CVD and cancer. There was a median lag of 5.5 years between the UK Biobank baseline when covariates measurements were taken and the accelerometry study, although covariates were stable over time, with the exception of medication [18]. In addition, adults' accelerometry-measured physical activity has been shown to be stable over time (for example, >90% of classification accuracy within one quartile over a period of 2-3 years) [42]. The responses to the baseline leisure time physical activity questions (including recreational walking) that formed the basis of our sample selection are subject to measurement error like any other self-reported measure and were also collected 5.5 years before the accelerometry study. However, the nonexerciser status among the UK Biobank accelerometry substudy participants with leisure time physical activity re-examination data was also stable over time (for example, 82%-88% retained the nonexerciser status). The UK Biobank had a very low response rate (5.5%) and it is not representative of the target population [43]. However, recent empirical work has shown that the poor representativeness of the UK Biobank sample does not materially influence the associations between physical activity and mortality outcomes [44].

In conclusion, we found that as few as two or three short bouts or approximately 3-4 min of VILPA per day were associated with substantially lower all-cause, CVD and cancer mortality risk. Although steeper mortality risk reductions occurred at the lower end of the VILPA distribution, there were continuing gains with larger amounts in a near-linear fashion. Individuals who find structured exercise unappealing or infeasible may consider exploring opportunities to introduce brief but regular bouts of VPA into their daily routines. VILPA in nonexercisers appears to elicit similarly beneficial associations with VPA in exercisers. Future guidelines could emphasize that potentially important health benefits could be accrued through VPA even among people who do not consider themselves to be formal "exercisers". Future trials and device-based cohort studies should further investigate the potential of VILPA (and any-domain VPA in general) as a time-efficient and potentially effective intervention for physically inactive and unfit adults. Our approach shows that wearable devices combined with machine learning-based methods and self-reported information can reveal physical activity "micro-patterns" as targets to prevent premature mortality, CVD and cancer in populations not willing and/or not able to engage in structured exercise during leisure time.

3. Methods

3.1. Sample and design

Figure 1 describes the derivation of the analytic sample. The UK Biobank Study is a prospective cohort study of adults aged between 40 and 69 years whose baseline measurements took place between 2006 and 2010. Participants provided informed consent and ethical approval was provided by the UK's National Health Service, National Research Ethics Service (Ethics Committee reference number: 11/NW/0382).

Between 2013 and 2015 (median 5.5 years after the baseline measurements), 103,684 UK Biobank participants wore a wrist-worn accelerometer for 7 days [24, 25]. We excluded participants with missing covariates and insufficient valid wear days. Monitoring days were considered valid if wear time was greater than 16 h. To be included in analysis, participants were required to have at least three valid monitoring days, with at least one of those days being a weekend day [45, 46]. We excluded participants who reported that they cannot walk.

To enable examination of VILPA in our study (brief bouts of nonex ercise VPA occurring during daily living), we included only participants who report-

ed no leisure time exercise participation and no more than one recreational walk per week. Participation in exercise and recreational walking was measured through a close-ended touch-screen questionnaire that asked participants to report if, how often, and for how long they participate in such activities (Supplementary Table 2). Among the included 14,982 participants who were walking for recreation once a week or less, the average spacing of VILPA bouts was 165.7 (47.0) min within days and 16.7 (5.5) h between days (last session of a day versus first session the day after). The modal median length of the (at most) one and only weekly walking session these participants reported was 30-60 min (32.5% of the 14,982 participants), effectively eliminating the possibility that the device-recorded VILPA bouts occurred during recreational walking.

To provide a comparison between effects of VILPA and (context-agnostic) VPA we repeated the main analyses among "exercisers", defined as those UK Biobank accelerometry substudy participants who did not meet the above criteria to be considered nonexercisers; that is, those who reported any leisure time exercise or more than one recreational walking session per week (Supplementary Table 1).

3.2. *Definition of VILPA and choice of bout length*

We based the choice of VILPA bout length entered in our analyses on an ongoing study of 58 adults (mean age 55.7 (s.d. 10.1) years) aimed at developing an empirical definition of VILPA (M.N.A., N. Johnson, C.T.-N., M.J.G. and E.S., unpublished data). Participants completed five activities of daily living while wearing an indirect calorimetry unit (Cosmed K5) and Polar heart-rate monitor. The activities included: (1) walking on a flat surface at a self-selected "very fast" pace; (2) walking on a flat surface while carrying shopping-like bags equivalent to 5% of body weight at a self-defined "fast" pace; (3) walking on a flat surface while carrying shopping-like bags equivalent to 10% of body weight at a self-defined "fast" pace; (4) walking at a 2.5% gradient at a self-defined "very fast" pace (treadmill); and (5) walking at a 7.0% gradient at a self-defined "very fast" pace (treadmill). The sequence of activities was randomized for each participant and counterbalanced across participants to prevent biases due to residual fatigue accumulating during the protocol.

Participants performed each activity until vigorous intensity was reached for two of three criteria: (1) %VO_2max (percentage of maximal oxygen updake) (≥64%); (2) %HRmax (percentage of maximal heart rate) (≥77%); and (3) rating of perceived exertion (Borg scale) ≥15. For %VO_2max and %HRmax, the threshold had to be met for at least 30 consecutive seconds to minimize the effects of noise. VO_2max was calculated using the Ebbeling treadmill test and HRmax was calculated using the Tanaka equation [47]. Between activities, participants had 5 min of seated recovery, or until heart rate and breathing returned to resting levels. Resting VO_2 and heart rate were measured at the beginning of each session with the participant lying supine using 5 min of steady-state (coefficient of variation ≤ 10%). The duration to reach vigorous intensity across all five activities is shown in Supplementary Table 7. As the mean time required to reach vigorous intensity in two of the above three physiological intensity indices was 73.5 s (s.d. 26.2 s) across all activities, we decided to test VILPA bouts lasting up to 1 and up to 2 min in the present analyses. As the length of raw bouts within these two VILPA frequency exposures was highly variable, we length-standardized analytic bouts to one minute (for raw bouts lasting up to 1 minute) or two minutes (for raw bouts lasting up to 2 minutes) using a rolling sum on the time-series data until 1 or 2 minutes, respectively, was reached or exceeded. For example, a participant with five consecutive raw bouts lasting up to 1 minute each (20, 30, 20, 40, and 10 seconds long), would be assigned 1.83 analytic bouts: the first three raw bouts would count as one and the rolling sum would be reset; then the last two raw counts would count as 0.83 length-standardised bouts (50 seconds divided by 60). This bout handling has analytic and interpretational advantages: a) it mitigates against the problem of multicollinearity between raw VILPA frequency and dai-

ly VILPA duration, and b) permits a more concrete behavioural interpretation of the VILPA frequency findings than raw bouts, as each length-standardised bout can be specifically interpreted as lasting 1 minute or 2 minutes.

3.3. *Wearable device-based physical activity classification*

The methods we describe here were used to classify physical activity intensity in both the nonexercisers (main analyses) and exercisers (additional analyses) strata. Supplementary Fig. 4 summarizes how activity intensity was classified using a previously validated random forest (RF) activity classifier [33]. RF is an ensemble of multiple decision trees. Each tree is learned on a bootstrap sample of training data and each node in the tree is split using the best among a randomly selected set of acceleration features. The decisions from each tree are aggregated and a final model prediction is based on majority vote. The RF model requires very little preprocessing of the data because the features do not need to be normalized. In addition, the model is resistant to over-fitting the training data because each tree within the forest is independently grown to maximum depth using a randomly selected subset of features.

This two-stage classifier first categorized physical activity in 10-s windows into one of four activity classes: sedentary, standing utilitarian movements (for example, ironing a shirt, washing dishes), walking activities (for example, gardening, active commuting, mopping floors), running/high energetic activities (for example, active playing with children). These activity classes were then assigned to one of four activity intensities: sedentary, light, moderate and vigorous. Walking activities were classified as light (an acceleration value of <100 mg), moderate (≥100 mg) and vigorous (≥400 mg) intensity [48]. For example, for a VILPA bout lasting up to 2 min, 12 consecutive 10-s windows needed to be classified as vigorous. When there were more than 12 consecutive vigorous activity windows, these bouts counted as long VPA sessions in the corresponding analyses (2.3% of all VPA bouts). Differentiation between sleep [36] and nonwear [35] was identified using the change in tilt angle and acceleration standard deviation. Monitors were calibrated [49] and corrected for orientation [50] using previously published methods, although residual signal and alignment uncertainties may persist.

Activities in an independent sample of 98 participants (age 56.4 ± 15.7 years ; 53.1% female) from the US [51] (University of California Irvine Center for Machine Learning and Intelligent Systems Physical Activity Monitoring for Aging People study (published data), accessible at https://archive.ics.uci.edu/ml/datasets) and Australia [52] (University of Queensland Where and When at Work study (published data) and University of Sydney Intermittent Lifestyle Physical Activity Study (unpublished data)) providing 103,607 activity samples from structured and free-living activities (17,267 min) were used to assess robustness and generalizability of the classifier (Supplementary Tables 8 and 9). For free-living activities participant-worn or researcher-held Go-Pro video-recordings were used to attain ground-truth physical activity. Video files were imported into the Noldus Observer XT software v16.0 for continuous direct observation coding. A two-stage direct observation scheme was implemented in which the participant's movement behavior was coded for activity type and then activity intensity based on the Compendium of Physical Activities [53]. The direct observation system generated a vector of date–time stamps corresponding to the start and finish of each movement event, which were used to assign the activity codes to the corresponding time segments of the accelerometer data. Interobserver reliability was assessed by dual coding. The intraclass correlation coefficient for coding activities was 0.91 (0.87-0.94).

Performance was further evaluated in a separate sample of 151 adults (age range 18-91 years, 65.6% female; Supplementary Fig. 5) recruited from the UK [34] (University of Oxford Capture 24 study (published data), accessible at https://ora.ox.ac.uk/objects/ uuid:99d7c092-d865-4a19-b096-cc16440cd001). Participants in this data set wore body cameras that provided pictures every 20 s to anno tate ground-truth free-living activity labels.

The picture-based activity coding scheme has been previously described [34]. A total of 172,360 activity samples (28,727 min) were provided by participants.

3.4. *Outcome ascertainment*

Because of the nature of rolling updates for the data linkage, participants were followed up to 31 October 2021, with deaths obtained through linkage with the National Health Service (NHS) Digital of England and Wales or the NHS Central Register and National Records of Scotland. CVD mortality was defined as death attributed to diseases of the circulatory system, excluding hypertension, diseases of arteries and lymph (ICD-10 codes: I0, I11, I13, I20-I51, I60-I69). Cancer mortality was defined as death attributed to any cancer excluded in situ, benign, uncertain, nonmelanoma skin cancer or non-well-defined cancers (ICD-10 codes beginning "C0", "C1", "C2", "C3", "C4" (excluding C49.9), "C5", "C6", "C70", "C71", "C72", "C73", "C74", "C75", "C7A", "C8" or "C9").

3.5. *Statistical analyses*

In our study, the range of VILPA values (and context-agnostic VPA values in exercisers) was capped at the 97.5 percentile to minimize the influence of sparse data. To reduce the possibility of reverse causation through prodromal/undiagnosed disease, all analyses excluded those with an event within the first 2 years of follow-up. We also excluded those with prevalent CVD and prevalent cancer at baseline (CVD and cancer mortality analyses, respectively).

We examined the dose–response of average daily duration and frequency of VILPA bouts lasting up to 1 min and up to 2 min using Cox proportional hazards (all-cause mortality) and Fine–Gray subdistribution hazards to account for competing mortality risks (CVD and cancer mortality) [54]. In all analyses, we set knots at the 10th, 50th and 90th percentiles. Departure from linearity was assessed by a Wald test. Proportional hazards assumptions were tested using Schoenfeld residuals in the models with all three outcomes and no violations were observed (all $P > 0.05$). Analyses were adjusted for age, sex, daily duration of light- and of moderate-intensity physical activity, mutual adjustment for daily duration and frequency of vigorous-intensity physical activity bouts lasting more than 1 to 2 min as appropriate, smoking, alcohol, accelerometry-estimated sleep duration [35, 36], fruit and vegetable consumption, education, parental history of CVD and cancer, medication use (insulin, blood pressure, cholesterol). All-cause mortality analyses were also adjusted for prevalent CVD and cancer, CVD analyses were adjusted for prevalent cancer, and cancer analyses were adjusted for prevalent CVD (Supplementary Table 3 provides full covariate definitions).

In the exercisers stratum of the UK Biobank accelerometry sample, we repeated the above multivariable-adjusted analyses for daily duration and frequency of (context-agnostic) VPA for bouts lasting up to 2 min, and we compared findings with the equivalent VILPA findings using overlay dose–response plots.

To assert the degree to which VILPA and VPA may contribute to mortality beyond the associations of overall movement volume, we also carried out a volume analysis based on energy expenditure using methods analogous to the study by Strain *et al.* [18] We calculated physical activity energy expenditure for all VILPA and VPA bouts lasting up to 2 min.

To provide conservative point estimates we calculated the "minimal dose", defined as VILPA volume/frequency associated with 50% of the optimal risk reduction [37, 38]. We also present point estimates (HRs and 95% CI) associated with the median and maximum volume/frequency VILPA values. We calculated E-values to estimate the plausibility of bias from unmeasured confounding [30, 55].

We conducted sensitivity analyses of VILPA with additional adjustment for body mass index. To investigate potential reverse causation bias we also excluded participants who had poor self-rated health. In another sensitivity analysis, we tested the influence of applying a conservative definition

of "nonexercisers" by restricting analyses to the 10,230 participants who reported no recreational walking and no leisure time exercise.

We performed all analysis using R statistical software v.4.2.1 with RMS v.6.3.0 and survival package v.3.3.1.

We reported this study as per the Strengthening the Reporting of Observational Studies in Epidemiology guidelines (Supplementary Table 10).

3.6. Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

3.7. Data availability

The UK Biobank data that support the findings of this study can be accessed by researchers on application (https://www.ukbiobank. ac.uk/register-apply/). Variables derived specifically for this study will be returned along with the code to the UK Biobank for future applicants to request.

Availability of other datasets related to the study: University of California Irvine Center for Machine Learning and Intelligent Systems Physical Activity Monitoring for Aging People: https://archive. ics. uci.edu/ml/datasets; University of Queensland Where and When at Work study: available upon reasonable request to the study's PI [11]; University of Sydney Intermittent Lifestyle Physical Activity Study: available upon reasonable request to the authors; University of Oxford Capture 24 study: https://ora.ox.ac.uk/objects/uuid:99d7c092-d865-4a19-b096-cc16440cd001

4. Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-022-02100-x.

5. Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 25813. The authors would like to thank all the participants and professionals contributing to the UK Biobank, and S. Paudel for her assistance with an early short report version of this manuscript. This study was funded by an Australian National Health and Medical Research Council Investigator Grant (APP1194510) and Ideas Grant (APP1180812). A.D. is supported by the Wellcome Trust (223100/Z/21/Z), National Institute for Health Research Oxford Biomedical Research Centre, Novo Nordisk, the British Heart Foundation Centre of Research Excellence (grant number RE/18/3/34214), the Alan Turing Institute and the British Heart Foundation (grant number SP/18/4/33803) and Health Data Research UK, an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities.

6. Author contributions

E.S. and M.N.A. had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. E.S. conceived the idea, designed the study and drafted the manuscript with the assistance of M.N.A., and obtained funding. M.N.A. processed and prepared the wearable sensors' data and carried out the statistical analyses of the paper. M.N.A., A.D. and J.M.R.G. revised the analytic design. A.D. critically evaluated the physical activity classification schema. All authors contributed to the interpretation of the data and provided critical revisions of the manuscript for important intellectual content over multiple rounds.

7. Code availability

The statistical code used in the analyses of this manuscript can be downloaded from https://zenodo.org/record/7187927#.Y0ZfoHZBy3A

8. References

1. Kraus, W. E. *et al.* Physical activity, all-cause and cardiovascular mortality, and cardiovascular disease. *Med. Sci. Sports Exerc.* **51**, 1270-1281 (2019).
2. Rezende, L. F. M. D. *et al.* Physical activity and cancer: an umbrella review of the literature including 22 major anatomical sites and 770 000 cancer cases. *Br. J. Sports Med.* **52**, 826-833 (2018).
3. Moore, S. C. *et al.* Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. *JAMA Intern. Med.* **176**, 816-825 (2016).
4. Bull, F. C. *et al.* World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* **54**, 1451-1462 (2020).
5. Piercy, K. L. *et al.* The physical activity guidelines for Americans. *JAMA* **320**, 2020-2028 (2018).
6. World Health Organization. *Global Recommendations on Physical Activity for Health* (World Health Organization, 2010).
7. U.S. Department of Health and Human Services. *2008 Physical Activity Guidelines for Americans* (U.S. Department of Health and Human Services, 2008).
8. Oja, P. *et al.* Associations of specific types of sports and exercise with all-cause and cardiovascular-disease mortality: a cohort study of 80 306 British adults. *Br. J. Sports Med.* **51**, 812-817 (2017).
9. Stamatakis, E. & Chaudhury, M. Temporal trends in adults' sports participation patterns in England between 1997 and 2006: the Health Survey for England. *Br. J. Sports Med.* **42**, 901-908 (2008).
10. Gebel, K. *et al.* Effect of moderate to vigorous physical activity on all-cause mortality in middle-aged and older Australians. *JAMA Intern. Med.* **175**, 970-977 (2015).
11. Wang, Y., Nie, J., Ferrari, G., Rey-Lopez, J. P. & Rezende, L. F. Association of physical activity intensity with mortality: a national cohort study of 403 681 US adults. *JAMA Intern.* Med. **181**, 203-211 (2021).
12. Rey Lopez, J. P., Gebel, K., Chia, D. & Stamatakis, E. Associations of vigorous physical activity with all-cause, cardiovascular and cancer mortality among 64 913 adults. *BMJ Open Sport Exerc. Med.* **5**, e000596 (2019).
13. Rey Lopez, J. P., Sabag, A., Martinez Juan, M., Rezende, L. F. M. & Pastor-Valero, M. Do vigorous-intensity and moderate-intensity physical activities reduce mortality to the same extent? A systematic review and meta-analysis. *BMJ Open Sport Exerc. Med.* **6**, e000775 (2020).
14. Ross, R. *et al.* Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association.*Circulation* **134**, e653-e699 (2016).
15. World Cancer Research Fund & American Institute for Cancer Research. Physical activity and the risk of cancer. *Continuous Update Project* https://www.wcrf.org/wp-content/ uploads/2021/02/ Physical-activity.pdf (2018).
16. Matthews, C. E. *et al.* Amount and intensity of leisure-time physical activity and lower cancer risk. *J. Clin. Oncol.* **38**, 686-697 (2020).
17. O'Donovan, G., Lee, I.-M., Hamer, M. & Stamatakis, E. Association of "weekend warrior" and other leisure time physical activity patterns with risks for all-cause, cardiovascular disease, and cancer mortality. *JAMA Intern. Med.* **177**, 335-342 (2017).
18. Strain, T. *et al.* Wearable-device-measured physical activity and future health risk. *Nat. Med* **26**, 1385-1391 (2020).
19. Stamatakis, E. *et al.* Untapping the health enhancing potential of vigorous intermittent lifestyle physical activity (VILPA): rationale, scoping review, and a 4-Pillar research framework. *Sports Med.* **51**, 1-10 (2021).
20. Rey-Lopez, J. P., Stamatakis, E., Mackey, M., Sesso, H. D. & Lee, I. M. Associations of self-reported stair climbing with all-cause and cardiovascular mortality: The Harvard Alumni Health Study. *Prev. Med. Rep.* **15**, 100938 (2019).
21. Celis-Morales, C. A. *et al.* The association between physical activity and risk of mortality is modulated by grip strength and cardiorespiratory fitness: evidence from 498 135 UK-Biobank participants. *Eur. Heart J.* **38**, 116-122 (2017).
22. Lamoureux, N. R. *et al.* Temporal trends in the cardiorespiratory fitness of 2,525,827 adults between 1967 and 2016: a systematic review. *Sports Med.* **49**, 41-55 (2019).
23. Bennie, J. A. *et al.* The descriptive epidemiology of total physical activity, muscle-strengthening exercises and sedentary behaviour among Australian adults – results from the National Nutrition and Physical Activity Survey. *BMC Public Health* **16**, 73 (2016).
24. Doherty, A. *et al.* Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank Study. *PLoS ONE* **12**, e0169649 (2017).
25. Ramakrishnan, R. *et al.* Accelerometer measured physical activity and the incidence of cardiovascular disease: evidence from the UK Biobank cohort study. *PLoS Med.* **18**, e1003487 (2021).
26. Stevens, M. L. *et al.* Thigh-worn accelerometry for measuring movement and posture across the 24-hour cycle: a scoping review and expert statement. *BMJ Open Sport Exerc. Med.* **6**, e000874 (2020).
27. Stamatakis, E. *et al.* Emerging collaborative research platforms for the next generation of physical activity, sleep and exercise medicine guidelines: the Prospective Physical Activity, Sitting, and Sleep Consortium (ProPASS). *Br. J. Sports Med.* **54**, 435-437 (2020).
28. Dowd, K. P. *et al.* A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. *Int. J. Behav. Nutr. Phys. Act.* **15**, 15 (2018).
29. Hiral, M. J. *et al.* Association of step counts over time with the risk of chronic disease in the All of Us Research Program. *Nat. Med.* **28**, 2301-2308 (2022).
30. Ahmadi, M. N. *et al.* Vigorous physical activity incident heart disease and cancer: how little is enough? *Eur. Heart J.* ehac572 https://doi.org/10.1093/eurheartj/ehac572 (2022).

31. Larrichia, F. Wearables – statistics & facts. *Statista* https://www.statista.com/topics/1556/wearable-technology/#topicHeaderwrapper (2022).
32. Walmsley, R. *et al.* Reallocation of time between device-measured movement behaviours and risk of incident cardiovascular disease. *Br. J. Sports Med.* **56**, 1008-1017 (2022).
33. Pavey, T. G., Gilson, N. D., Gomersall, S. R., Clark, B. & Trost, S. G. Field evaluation of a random forest activity classifier for wrist-worn accelerometer data. *J. Sci. Med. Sport* **20**, 75-80 (2017).
34. Willetts, M., Hollowell, S., Aslett, L., Holmes, C. & Doherty, Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci. Rep.* **8**, 7961 (2018).
35. Ahmadi, M. N., Nathan, N., Sutherland, R., Wolfenden, L. & Trost, S. G. Non-wear or sleep? Evaluation of five non-wear detection algorithms for raw accelerometer data. *J. Sports Sci.* **38**, 399-404 (2020).
36. van Hees, V. T. *et al.* Estimating sleep parameters using an accelerometer without sleep diary. *Sci. Rep.* **8**, 12975 (2018).
37. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose-response analysis using R. *PLoS ONE* **10**, e0146021 (2015).
38. Rampinelli, C. *et al.* Exposure to low dose computed tomography for lung cancer screening and risk of cancer: secondary analysis of trial data and risk–benefit analysis. *BMJ* **356**, j347 (2017).
39. Allison, M. K. *et al.* Brief intense stair climbing improves cardiorespiratory fitness. *Med. Sci. Sports Exerc.* **49**, 298-307 (2017).
40. Batacan, R. B., Duncan, M. J., Dalbo, V. J., Tucker, P. S. & Fenning, S. Effects of high-intensity interval training on cardiometabolic health: a systematic review and meta-analysis of intervention studies. *Br. J. Sports Med.* **51**, 494-503 (2017).
41. Han, M. *et al.* Cardiorespiratory fitness and mortality from all causes, cardiovascular disease and cancer: dose–response meta-analysis of cohort studies. *Br. J. Sports Med.* **56**, 733-739 (2022).
42. Keadle, S. K. *et al.* Reproducibility of accelerometer-assessed physical activity and sedentary time. *Am. J. Prev. Med.* **52**, 541-548 (2017).
43. Fry, A. *et al.* Comparison of sociodemographic and health- related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026-1034 (2017).
44. Stamatakis, E. *et al.* Is cohort representativeness passé? Poststratified associations of lifestyle risk factors with mortality in the UK Biobank. *Epidemiology* **32**, 179-188 (2021).
45. Del Pozo Cruz, B., Ahmadi, M.N., Lee, I.M. & Stamatakis, E. Prospective associations of daily step counts and intensity with cancer and cardiovascular disease incidence and mortality and all-cause mortality. *AMA Intern. Med.* **182**, 1139-1148 (2022).
46. Del Pozo Cruz, B., Ahmadi, M., Naismith, S.L. & Stamatakis, E. Association of daily step count and intensity with incident dementia in 78 430 adults living in the UK. *JAMA Neurol.* **79**, 1059-1063 (2022).
47. Tanaka, H., Monahan, K. D. & Seals, D. R. Age-predicted maximal heart rate revisited. *J. Am. Coll. Cardiol.* **37**, 153-156 (2001).
48. Hildebrand, M., VT, V. A. N. H., Hansen, B. H. & Ekelund, U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med. Sci. Sports Exerc.* **46**, 1816-1824 (2014).
49. Sipos, M., Paces, P., Rohac, J. & Novacek, P. Analyses of triaxial accelerometer calibration algorithms. *IEEE Sens. J.* **12**, 1157-1165 (2011).
50. Mizell, D. Using gravity to estimate accelerometer orientation. In *Proc. Seventh IEEE International Symposium on Wearable Computers* 252 (Citeseer, 2003).
51. Reiss, A., Weber, M. & Stricker, D. Exploring and extending the boundaries of physical activity recognition. In *2011 IEEE International Conference on Systems, Man, and Cybernetics* 46-50 (IEEE, 2011).
52. Clark, B. K., Winkler, E. A., Brakenridge, C. L., Trost, S. G. & Healy,G. N. Using Bluetooth proximity sensing to determine where office workers spend time at work. *PLoS ONE* **13**, e0193971 (2018).
53. Ainsworth, B. *et al.* Second update of codes and MET values. *Med. Sci. Sports Exerc.* **39**, 1575-1581 (2011).
54. Austin, P. C. & Fine, J. P. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat. Med.* **36**, 4391-4400 (2017).
55. Haneuse, S., VanderWeele, T.J. & Arterburn, D. Using the E-value to assess the potential effect of unmeasured confounding in observational studies. *JAMA* **321**, 602-603 (2019).

Mikropoluitzaileak araztegietan pandemiaren ondoren

COVID-19aren itxialdian, asko handitu zen botika terapeutikoen erabilera, birusaren aurkako tratamendu berezirik ez zegoelako. Botika horiek, askotan, uretara joaten dira, eta hondakin-uren tratamendu-zentroetan azaltzen dira, uraren analisi kimikoan. Substantzia horiek berriak dira araztegietan, eta metodo berriak bilatu beharko dira urak garbitzeko. Arazte-teknika berriak eskuragarri daudenean, haiek inplementatzeko baliabideak jarri beharko dira.

Pandemian zehar, Europan araztegietako urak analizatu izan ziren, bai birusaren hedapen geografikoari jarraitzeko, bai eta ustezko mikropoluitzaile berriak bilatzeko ere. Euskadin ere egin zuten; ikertzaile-talde batek bi araztegitako urak analizatu zituen, Crispijanakoa Araban eta Galindokoa Bizkaian, 2020ko apiriletik uztailera bitartean.

Zehaztasun handiko analisien bitartez, substantzia askoren aztarnak detektatu zituzten, hondakin-urak tratatzeko instalazioak ez baitziren eraginkorrak substantzia horiek desagerrarazteko. Kasu batzuetan, substantzien kontzentrazioa ohi baino handiagoa zen; medikamentu antibiralena eta antimikrobianoena, adibidez. Gainera, analisi horien emaitzek adierazten dute antsiolitikoen erabilerak gora egin zuela, bai eta legez kontrako zenbait drogarenak ere. Egoera berriak substantzia berriak ekarri ditu araztegietara, eta arriskua dago substantzia horiek ez detektatzeko. Beraz, analisi-teknikak berak ere aztertu behar izan dira, ahalik eta mikropoluitzaile gehien atzeman ahal izateko.

Pandemiaren eragin kimikoa

Crispijanako eta Galindoko araztegietan, ur-laginak hartu, iragazi eta molekulak oso zehaztasun handiz identifikatzen dituen teknika batez aztertu ziren: substantziak likido-kromatografo batez banatu eta gero, banaka identifikatu ziren masa-espektroskopiaren bitartez.

Bi araztegietan antzeko substantziak topatu dituzte. Gehienbat, produktu farmazeutikoak dira, baina bestelako substantzia asko ere badaude: estimulatzaileak, pestizidak, hormonak, produktu industrialak, suaren kontrako produktuak eta abar. Ohiko medikamentu batzuk ohi baino kontzentrazio altuagoetan azaldu dira, azetaminofenoa (parazetamola), metformina (diabetesaren kontrako botikarik ohikoena) eta kafeina, adibidez. Beste batzuk lehen aldiz azaldu dira Euskadiko araztegi batean. Arreta berezia jaso du COVID-19a tratatzeko ahaleginen ondorioak. Ustezko tratamendu berriei lotutako substantziak agertu dira lehen aldiz araztegietan, hidroxiklorokina eta lopinavir esate baterako. Gainera, antipsikotikoen kopuruak ere handitu dira, eta gauza bera gertatu da legez kanpoko droga batzuekin ere, anfetaminarekin eta ketaminarekin adibidez.

«Araztegietan botiken aztarnak areagotu dira, eta arriskua dago substantzia horiek ez detektatzeko»

Aztertu behar da substantzia horiek ingurumenean kalte egiten ote duten; alegia, haien biotoxikotasuna neurtzeko irizpideak ezarri behar dira. Oro har, molekula bakoitzak berezko inpaktua du.

Araztegietan, helburua da substantzia horiek uretatik kentzea. Hain zuzen ere, arazte-teknikak garatu, eta etorkizunean teknika horiek inplementatzeko baliabideak jarri beharko dira.

InpAkta 00 (2023) 125-149

http://dx.doi.org/10.1016/j.scitotenv.2023.162281

Comprehensive micropollutant characterization of wastewater during Covid-19 crisis in 2020: suspect screening and environmental risk prioritization strategy

N. Lopez-Herguedas[1,2], M. Irazola[1,2], I. Alvarez-Mora[1,2], G. Orive[3,4], U. Lertxundi[4,5], M. Olivares[1,2], O. Zuloaga[1,2], A. Prieto[1,2]

[1] Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country (UPV/EHU); [2] Research Centre for Experimental Marine Biology and Biotechnology (PIE), University of the Basque Country (UPV/EHU); [3] NanoBioCel Group, Laboratory of Pharmaceutics, Faculty of Pharmacy, University of the Basque Country UPV/EHU; [4] Bioaraba, NanoBioCel Research Group; [5] Bioaraba Health Research Institute; Osakidetza Basque Health Service, Araba Mental Health Network, Araba Psychiatric Hospital, Pharmacy Service

ABSTRACT: Micropollutants monitoring in wastewater can serve as a picture of what is consuming society and how it can impact the aquatic environment. In this work, a suspect screening approach was used to detect the known and unknown contaminants in wastewater samples collected from two wastewater treatment plants (WWTPs) located in the Basque Country (Crispijana in Alava, and Galindo in Vizcaya) during two weekly sampling campaigns, which included the months from April to July 2020, part of the confinement period caused by COVID-19. To that aim, high-resolution mass spectrometry was used to collect full-scan data-dependent tandem mass spectra from the water samples using a suspect database containing >40000 chemical substances. The presence of more than 80 contaminants was confirmed (level 1) and quantified in both WWTP samples, while at least 47 compounds were tentatively identified (2a). Among the contaminants of concern, an increase in the occurrence of some compounds used for COVID-19 disease treatment, such as lopinavir and hydroxychloroquine, was observed during the lockdown. A prioritization strategy for environmental risk assessment was carried out considering only the compounds quantified in the effluents of Crispijana and Galindo WWTPs. The compounds were scored based on the removal efficiency, estimated persistency, bioconcentration factor, mobility, toxicity potential and frequency of detection in the samples. With this approach, 33 compounds (e.g. amantadine, clozapine or lopinavir) were found to be considered key contaminants in the analyzed samples based on their concentration, occurrence and potential toxicity. Additionally, antimicrobial (RQ-AR) and antiviral (EDRP) risk of certain compounds was evaluated, where ciprofloxacin and fluconazole represented medium risk for antibiotic resistance (1>RQ-AR>0.1) in the aquatic ecosystems. Regarding mixture toxicity, the computed sum of toxic unit values of the different effluents (>1) suggest that interactions between the compounds need to be considered for future environmental risk assessments.

* **Corresponding author / Harremanetan jartzeko:** N. Lopez-Herguedas, Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country UPV/EHU, Leioa, Basque Country, Spain. e-mail: naroa.lopez@ehu.eus. **Contributed by** Jon Zarate **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Lopez-Herguedas, N. *et al.* «Comprehensive micropollutant characterization of wastewater during Covid-19 crisis in 2020: suspect screening and environmental risk prioritization strategy», *Science of the Total Environment*, 873 (2023) 162281 (http://dx.doi.org/10.1016/j.scitotenv.2023.162281)

1. Introduction

The year 2020 was marked by the onset of the global pandemic triggered by the SARS-CoV-2 virus, causing millions of deaths all over the world (WHO, 2021). This situation led most of the countries to introduce several measures (e.g. cancellation of public events, closure of schools and various businesses, curfews and lockdowns) in order to avoid the spread of the virus. This standstill of the countries severely affected the health, the economy and the social life of most of citizens all over the world.

During the pandemic period wastewater was used in many research studies to monitor the spread of the virus considering its excretion from infected people (de Araújo *et al.*, 2022; Godini *et al.*, 2021; Kuroda *et al.*, 2021) but also to determine whether people lifestyle changed. In fact, the lack of specific therapeutic treatments to combat COVID-19 led to an unprecedented consumption of different therapeutic drugs (Cappelli *et al.*, 2022; Kuroda *et al.*, 2021), which could end-up in environmental waters (Bandala *et al.*, 2021; Cappelli *et al.*, 2022; Domingo-Echaburu *et al.*, 2022). Particularly, during the confinement time, high amounts of antiviral and/or antimicrobial pharmaceuticals were prescribed for COVID-19 treatment and their inefficient elimination in wastewater treatment plants (WWTPs) led to detect such compounds in wastewater effluents and environmental waters (Nannou *et al.*, 2020). Moreover, the potential presence of antivirals and antimicrobials in environmental waters may increase the development of antiviral (Kuroda *et al.*, 2021; Nannou *et al.*, 2020) and antimicrobial resistance (Knight *et al.*, 2021; Usman *et al.*, 2020). In this regard, it is known that the environment constitutes one of the main sources of gene resistance to pathogens (Bengtsson-Palme and Larsson, 2016), but such resistance is not considered in the current regulatory systems (Boxall *et al.*, 2012). Even though efforts have been done to determine the minimum inhibitory concentration (MIC) of certain compounds with antimicrobial activity (Bengtsson-Palme and Larsson, 2016; Booth *et al.*, 2020), adverse effects even below the MIC values have been reported in the literature (Andersson and Hughes, 2012; Gullberg *et al.*, 2014), pointing out the lack of comprehensive knowledge about the effects of the unknown chemicals' cocktail can pose on the environment and human health (Fonseca *et al.*, 2020; Markert *et al.*, 2020; Nilsen *et al.*, 2019).

The potential of wastewater monitoring to get epidemiological information on human consumption and exposure to chemical residues has been widely demonstrated in many research works, where wastewater-based epidemiology (WBE) approach was used (Alygizakis *et al.*, 2021; Been *et al.*, 2021; Galani *et al.*, 2021; Nason *et al.*, 2022; Perkons *et al.*, 2022; Reinstadler *et al.*, 2021; Wang *et al.*, 2020). By monitoring wastewater samples during the pandemic period, for example, variations in benzoylecgonine use in European countries (Been *et al.*, 2021), increase of methamphetamine consumption (Reinstadler *et al.*, 2021), increase of benzodiazepines (psychoactive pharmaceuticals with anxiolytic activity) use (Alygizakis *et al.*, 2021) and no-alteration of certain pharmaceuticals consumption (Wang *et al.*, 2020) was determined using WBE approach.

As far as Spain is concerned, the monitoring of emerging contaminants (ECs) in wastewaters of WWTPs is widely done using mainly multi-targeted analytical methods (Afonso-Olivares *et al.*, 2017; Díaz-Garduño *et al.*, 2017; Martín *et al.*, 2012; Solaun *et al.*, 2021) and also applying WBE approach (Bijlsma *et al.*, 2021; Estévez-Danta *et al.*, 2022; Montes *et al.*, 2020). Although the unquestionable adequacy of target screening for the monitoring of a fixed set of micropollutants, the unknowns that may occur in the aquatic environment depends on many factors (e.g., land use, proximity to industry, type of sewer system, WWTP processes, population demographics, etc.) and contaminants end up being overlooked. Those limitations move scientists towards the use of more flexible and easily adaptable suspect screening studies that allow (i) addressing a larger amount of micropollutants and/or (ii) performing risk assessment (Cappelli *et al.*, 2022; Gago-Ferrero *et al.*, 2016; González-Gaya *et al.*, 2021; Li *et al.*, 2018; Perkons *et al.*, 2022). The use of those

analytical strategies to analyze wastewater samples can serve to determine as many as possible unknown chemicals which could provide hint information about what the population is consuming in a specific period of time.

Within this context, the main aim of this work was to evaluate the presence of micropollutants via suspect screening, and the subsequent confirmation through a validated target analysis in the influents and effluents of two WWTPs located in the Basque Country (Crispijana, Alava, and Galindo, Vizcaya) during two weekly sampling campaigns (from April to July 2020), in part of the period of confinement caused by COVID-19. The identification of the main potential toxicity drivers based on a prioritization strategy including different categories was assessed. Moreover, antimicrobial and antiviral compounds risks were also evaluated.

2. Experimental section

2.1. *Reagents and materials*

All chemicals and laboratory materials used in this work are provided in section S1 and the Supporting Information (SI) of Lopez-Herguedas *et al.* (Lopez-Herguedas *et al.*, 2022).

2.2. *Sampling*

Sampling was carried out 1 or 2 times per week, from April to July 2020 (Figure 1), collecting 24-hour composite aqueous samples (influent and effluents) from two WWTPs located in Vizcaya and Alava, Galindo and Crispijana, respectively (see details in section S2 in SI). Samples began to be collected after the peak incidence of Covid-19 cases in the Basque Country (Spain).

At the Galindo WWTP, composite samples were collected from the influent (IWW), primary treatment (EWW1), secondary treatment (EWW2) and tertiary treatment (EWW3), while at the Crispijana WWTP, the influent (IWW) and effluent after secondary treatment (EWW) were collected. All samples were stored and frozen at −20 °C until their analysis, which was carried out 2 months after their collection.

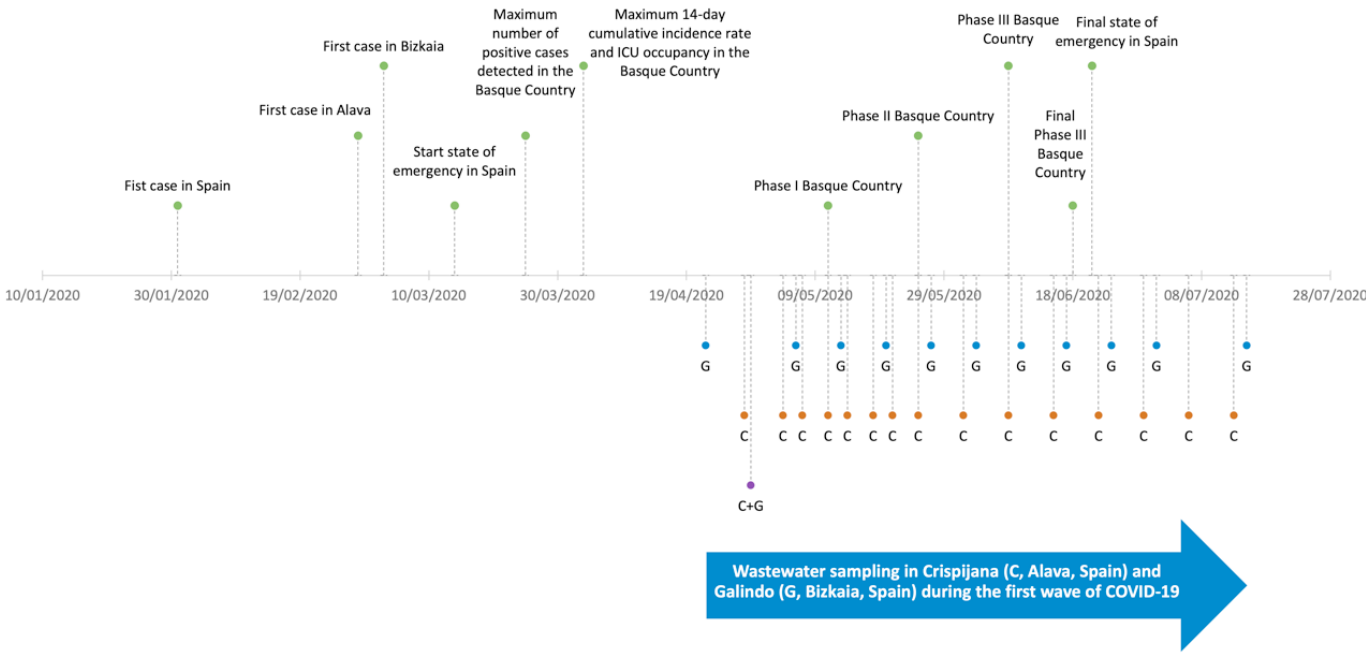


Figure 1. Timeline of Covid-19 situation in its first wave and sampling dates of composite water samples in both WWTPs (G: Galindo, C: Crispijana)

2.3. *Sample treatment*

The water samples were thawed and once at room temperature, all samples were filtered through cellulose filters (0.7 µm, 90 mm, Whatman; Maidstone, UK). Three replicates of 250 mL (effluent) or 100 mL (influent) were processed according to a previously validated method in our research group (González-Gaya *et al.*, 2021) (see details in section S2 in SI). Briefly, the samples were extracted with 500 mg solid-phase extraction (SPE) cartridges consisting of cation exchange (100 mg, ZT-WCX), anion exchange (100 mg, ZT-WAX) and reverse phase (300 mg, HRX) sorbents for effluent samples, and with 250 mg SPE cartridges containing half of the above-described amounts for influent samples. The cartridges were conditioned using 5 mL of MeOH:EtOAc and 5 mL of Milli-Q water. Subsequently, each sample aliquot was loaded and were left to dry under vacuum before analytes elution using 12 mL of a MeOH:EtOAc mixture (1:1) containing 2% ammonia and 12 mL of a MeOH:EtOAc mixture with 1.7% formic acid. Both extracts were combined, evaporated to dryness using a Turbovap (Zymark, Hopkinton, USA) under a gentle nitrogen stream and reconstituted in 250 µL of MeOH:Milli-Q water (1:1, v:v).

2.4. *Analysis by UHPLC-q-Orbitrap*

Extracts were analyzed on a Thermo Scientific Dionex Ulti-Mate 3000 UHPLC coupled to a Thermo Scientific Q Exactive Focus quadrupole-Orbitrap mass spectrometer (UHPLC-q-Orbitrap) equipped with a heated electrospray ionisation source (HESI, Thermo-Fisher Scientific, CA, USA) based on the previously developed methods (González-Gaya *et al.*, 2021; Lopez-Herguedas *et al.*, 2022) detailed in section S3 of the SI.

2.5. *Quality assurance of the analytical method*

The analytical protocol used in this work was thoroughly optimized in a previous work of our research group and is described elsewhere (González-Gaya *et al.*, 2021) (see section S4 in SI). Anyhow the QA/QC criteria of the analyses conducted in this work were assured for 231 compounds in terms of identification limits and apparent recoveries (see Table S1).

2.6. *Suspect analysis*

Suspect analysis data treatment was carried out using the Compound Discoverer 3.2 (Thermo-Fisher Scientific) and the workflow previously reported by González-Gaya *et al.* (González-Gaya *et al.*, 2021) (see detailed information in SI). Only Lorentzian peaks were considered and they were manually checked. The SusDat NORMAN database (40,059 compounds, www.norman-network.net, DOI:10.5281/zenodo.2664077) was used as a suspect list with a fixed error lower than ± 5 ppm in the exact mass. The molecular formulas suggested by the software were only accounted for if MS1 was satisfactorily matched (SFit> 30% and isotopic profile > 70%). Minimum peak areas considered were set at 10e6 units for both negative and positive ionization modes. Additionally, only peaks 10 times larger in the samples than in the blanks and with a relative standard deviation (% RSD) lower than 30% within injection replicates (n=3) were further studied. MS2 spectra were compared with mzCloud database (https://www.mzcloud.org/), and a match of over 70% was set for the identification of the feature. When the standards of the candidates were available, experimental retention time was confirmed with an allowed error of ± 0.1 min. If not available, retention times were estimated from the Retention Time Index (RTI) platform (http://rti.chem.uoa.gr/) and candidates were rejected or accepted depending on whether there was a statistical difference or not with the estimated value within the uncertainty of the model built. Finally, identification criteria according to Schymanski and coworkers (Schymanski *et al.*, 2014) was noted to provide the candidates with a tentative code from 1 to 3 levels of identification. Although this scale is numbered from one to five, in this work we annotated compounds up to level 3 being level 1 the one with the highest confidence level (features with their structure identified and confirmed by reference standard acquisition) and three the least one (features identified as potential

candidates with known structure but more than one candidate is provided since they are potential isomers).

2.7. *Quantification and multivariate data analysis*

Quantitative data analysis of the suspects annotated as level 1 (target analysis) was performed using Tracefinder 4.2 software (Thermo-Fisher Scientific). Target compounds and their instrumental characteristics including molecular formula, ionization mode, retention time (Rt) and experimental MS/MS fragments were added to the software library according to studies previously performed by the research group (Lopez-Herguedas *et al.*, 2022). To avoid false positives, the experimental retention time window was limited to 60 seconds around the retention time of the pure standard, a mass error equal to or less than 5 ppm, isotopic profile matching at more than 70% and mass accuracy for fragments equal to or less than 5 ppm were considered. Peak integration and calibration curves were checked manually.

Once obtained the data, principal component analysis (PCA) was carried out with PLS toolbox (8.7.1 version, Eigenvector Research, Wenatchee, USA) in the Matlab programming environment (R2019b, Mathworks Inc., Natick, USA). Mean-centering and variance scaling was carried out prior to multivariate statistical analysis. Leave-one-patient-out cross-validation was used to validate and optimize the PCA model.

2.8. *Prioritization strategy for environmental risk assessment*

Risk assessment was accomplished through a prioritization strategy of suspects annotated as level 1 following the approach described by Gros *et al.* with slight modifications (Gros *et al.*, 2017). Six category classes were set to prioritize the most environmentally relevant compounds identified in each WWTP effluent including: (a) removal efficiency (RE, %), (b) estimated persistency (half-life time in days, DT50), (c) bioconcentration factor (BCF), (d) mobility, (e) toxicity potential and (f) frequency of detection in the samples (Table 1). Each micropollutant was scored with a value between 1-5 in each category (a-e) summed up to obtain a total score, being the compounds showing the lowest value the ones posing the highest environmental risk. Compounds that were never detected above the LOQ were excluded in order to avoid overestimation of risks by including compounds that were likely to be absent. Similarly, compounds present at levels < LOQs in the influent samples were not considered since the calculated RE would be biased leading to an overestimation of the risk.

Table 1
Criteria and scoring system for prioritization of identified micropollutants

Criteria	Score				
	1	2	3	4	5
Removal efficiency (RE)	<40%		40-60%		>60%
Biodegradation (predicted half-life time in days)	>180	>60	>37.5	>15	<15
Bioaccumulation (BCFpred)	>10,000	>1000	>100	>10	<10
Mobility (log K_{ow})	<2.5		2.5-4.0		>4.0
Risk Quotient (RQ)	>1	>0.1	>0.01	>0.001	<0.001
Frequency of detection (%) in effluent	100%	>75%	>50%	>25%	<25%

RE (%) of individual ECs was estimated considering their concentrations in wastewater before and after wastewater treatment (Golovko *et al.*, 2021; Li *et al.*, 2018) (see **Equation 1**). Independent two samples t-test was performed at a 95% confidence level to evaluate significant differences among the concentrations quantified in influent and effluent samples for each contaminant to avoid comparison between influent and effluent pairs that do not really show significant differences and their comparison may lead to misleading results. Considering the high variability of the observed values between days, the scoring system for the RE relied on 3 values that were established as follows: (i) effectively removed compounds with RE values higher than 60%, (ii) moderately removed compounds with RE values between 40% and 60%, and (iii) not eliminated compounds with RE values lower than 40% and/or compounds for which influent and effluent mean concentrations are indistinguishable (e.g. DEP shows a RE of 65% in Galindo WWTP but the t-test reveals that values in the IWW and EWW3 are not significantly different).

$$RE(\%) = \left(\frac{\left([Influent] - [Effluent] \right)}{[Influent]} \right) \times 100 \qquad (1)$$

The biodegradation potential (due to biological activity, chemical reactivity or physical degradation) of the compounds is a good indicator of their persistence in the environment. The bioaccumulation potential refers to the ability that some chemical compounds have to accumulate in a living organism and can be predicted by the lipophilicity of the chemical. The values for both categories were defined based on Gros *et al.* (Gros *et al.*, 2017), which were established according to the European legislation for chemicals of concern, REACH (EC 1907/2006). In the present work, half-life times (DT50) and BCFs were retrieved from the CompTox Chemical Dashboard (https://comptox.epa.gov/dashboard/) relying on the OPERA models (Finckh *et al.*, 2022; Mansouri *et al.*, 2018).

The capability of a compound to diffuse the source to other environmental compartments is given by its mobility. Considering that log K_{ow} serves as a measure of the relationship between lipophilicity (fat solubility) and hydrophilicity (water solubility) of a substance, it was used to score the mobility pattern of compounds using the following criteria: (i) compounds with log K_{ow} < 2.5 were considered to be highly mobile, (ii) compound with log K_{ow} values between 2.5-4.0 were considered to show medium mobility, and compounds with log K_{ow} > 4.0 were considered to be low mobile (Dimitrov *et al.*, 2019; Jones-Lepp and Stevens, 2007; Roveri and Lopes Guimarães, 2023).

The toxicity potential was expressed in terms of risk quotients (RQ), calculated for each compound according to the European Union technical Guidance Document (European Parliament, 2006) as the ratio of the measured environmental concentration (MEC) in WWTP effluents and predicted no-effect concentration (PNEC). 95th percentiles of the measured concentrations for each compound were used as MEC values. The PNEC values were calculated as described by Lopez-Herguedas *et al.* (Lopez-Herguedas *et al.*, 2022) (see details in section S5 in SI).

Considering the sudden increase in the discharge of antimicrobials, including antibiotics and antivirals, to the environment the potential risk of the mentioned compounds was also determined. The Antibiotic Resistance (AR) was assessed based on the RQ metric (RQ-AR) as described by Bengtsson-Palme and Larsson (Bengtsson-Palme and Larsson, 2016). The PNECs for the selection of AR (PNEC-AR) were derived considering the MICs of the antibiotic compounds, which are the lowest concentrations of antibiotic for inhibiting bacterial growth, and the application of an appropriate assessment factor to the MIC (Bengtsson-Palme and Larsson, 2016; Cappelli *et al.*, 2022). On the other hand, the antiviral resistance was determined by the calculation of the Environmentally acquired antiviral Drug Resistance Potential (EDRP) as described by Kuroda and coworkers (Kuroda *et al.*, 2021) (Equation 2):

$$EDRP = Min\left(\frac{\mathrm{MEC\ 95^{th}\ perc}}{\mathrm{vEC_{50}\ or\ vIC_{50}}}, \frac{\mathrm{vEC_{50}\ or\ vIC_{50}}}{\mathrm{MEC\ 95^{th}\ perc}} \right) \qquad (2)$$

Where, vIC_{50} and vEC_{50} refer to the antiviral drug concentration which determines the 50% of the vi-

ral growth inhibition expressed as the half maximal inhibitory (IC_{50}) and effective (EC_{50}) concentrations, respectively. Those values were compiled from (Kuroda *et al.*, 2021). EDRP values vary between 0 and 1, being a value equal to 1 the maximum risk potential.

Given that the environmental samples are constituted by myriads of contaminants, mixture toxicity was also evaluated using the sum of toxic units (STU) approach based on CA (representing the worst-case scenario) in order to avoid an overestimation of the real risk as suggested by Backhaus and Faust, 2012 (Backhaus and Faust, 2012) (Equation 3):

$$RQ_{STU} = max\left(STU_{algae}, STU_{daphnids}, STU_{fish}\right) \times AF = max\left(\sum_{i=1}^{n}\frac{MEC}{EC50_{i,algae}}, \sum_{i=1}^{n}\frac{MEC}{EC50_{i,daphnids}}, \sum_{i=1}^{n}\frac{MEC}{EC50_{i,fish}}\right) \times AF \quad (3)$$

In this study, more conservative NOEC values corresponding to selected BQE instead of EC_{50} values were considered as reference concentrations for the calculation of STU to assess the impact on the aquatic ecosystem likewise for the calculation of individual RQ values. When experimental chronic NOEC values were not available, EC_{50} experimental values prevail over predicted NOEC values. In each case, an appropriate AF was applied (see section S5 in SI).

A dilution factor (DF) was applied to effluent concentrations to perform a more representative risk assessment caused by chemical exposure (Keller *et al.*, 2014). In both WWTPs, a minimum DF value was applied to simulate "the worst-case scenario"; thus, 10- and 50-fold effluent dilutions were considered for Crispijana and Galindo WWTP, respectively.

3. Results and discussion

The observations obtained in this work were based on a three-step workflow. First, the samples were analyzed using a suspect screening approach in order to detect the largest amount of contaminants present. Then, those candidates annotated as level 1 (i.e., standards available in the lab) were quantified. To end, those chemicals detected in secondary and tertiary effluent samples were ranked according to their potential hazards based on a prioritization strategy that included six relevant categories (see section 2.8).

3.1. Occurrence of ECs in analyzed samples

3.1.1. Suspect screening

The compounds identified and annotated at levels 1-3 by means of the workflow previously described (see section 2.6) are included in Table 2, where complete information about the annotation as well as the occurrence is compiled. In the case of Crispijana WWTP, among the identified candidates, the presence of 79 compounds was confirmed by chemical standards (level 1) (see section 3.2.1. and Table 2), while additionally, 47 candidates were tentatively identified as probable structures (level 2a) (29 candidates in IWW and 18 in EWW), and 4 tentative candidates (level 3) (only in IWW). Among the vast number of candidates identified some compounds stood out as the most frequently identified in Crispijana WWTP: (i) the pharmaceuticals lidocaine (anaesthetic), carbamazepine (anticonvulsant) and tramadol (analgesic) identified at level 1, and febuxostat (uric acid lowering agent) and rosuvastatin (antilipidemic) identified at level 2a; (ii) some transformation products identified at level 2a such as O-desmethylnaproxen, carbamazepine 10,11-epoxide and 11-ketotestosterone; and (iii) illicit drugs identified at level 2a such as ketamine and cocaine. Overall, more compounds with higher chromatographic areas were identified in influent wastewater, pointing out that the treatments implemented at the WWTPs partially removed chemicals present in wastewater.

Regarding the wastewaters from Galindo WWTP (see section 3.2.2. and Table 2), a total of 88 compounds were annotated as level 1, 53 candidates were annotated as level 2a (29 of them in the set of IWW and EWW1, 9 in the EWW2 and the remaining 15 in the EWW3), and 12 candidates (9 in the set of IWW and EWW1, 1 in the EWW2 and 2 in the EWW3) were tentatively identified (level 3). Compared to Crispijana WWTP, an increase in the number of identified compounds and chromatographic

areas was observed in the Galindo WWTP, a fact that may be related to the location (i.e. more populated area) and the influent volume (i.e., Galindo WWTP treats almost twice the flow that Crispijana WWTP treats). This is the case, for example, of methylparaben, nonylphenol, pyrantel or finasteride; compounds that were not identified in any sample from the Crispijana WWTP, but most of which were found in all influent samples belonging to Galindo. On the other hand, the tendency to find higher signals in IWW samples compared to the treated ones (EWW1, EWW2 and EWW3) remained constant, suggesting again a certain removal efficiency of the treatments implemented in the WWTPs.

3.1.2. Quantification of compounds annotated as level 1

The suspects annotated as level 1 were quantified using the chemical standards and following the QA/QC criteria described in section 2.5. The concentrations in ng/L found in all the studied samples (n = 32 and n = 47, in Crispijana and Galindo WWTPs, respectively) are detailed in Table 3 (see Tables S2 and S3 in SI for more detailed information). Multivariate data analysis was performed by means of PCA aiming to detect differences among the WWTPs studied as well as the different effluent treatments (see section S6 and Figure S1 in SI).

Among all the wastewater samples belonging to Crispijana WWTP, 80 compounds were quantified at ng/L level, whereas, 88 were the total compounds quantified in Galindo WWTP.

Overall, pharmaceutical products (PPs), stimulants, pesticides, phthalates, hormones, industrial agents, perfluorinated compounds and flame retardants were quantified at ng/L levels in both untreated and treated samples (i.e. IWW and EWW regarding Crispijana WWTP, IWW, EWW1, EWW2 and EWW3 regarding Galindo WWTP), being in both WWTPs the group of PPs the most abundant (around 59% and 65% of the detected compounds, respectively) (see Tables S2 and S3 in SI). Moreover, as it is summarized in Table 3, most of the compounds detected in Crispijana WWTP were also detected in Galindo WWTP. Following the trend observed in suspect screening, the highest concentration levels were found in IWW samples suggesting the removal efficiency of the treatments for some of the detected compounds. Concretely, the pharmaceuticals acetaminophen, (also known as paracetamol, an anti-inflammatory used to treat headaches), metformin (a drug to treat diabetes) and mycophenolic acid (an antibiotic usually used as an immunosuppressant drug, in organ transplants or for the treatment of certain autoimmune diseases), as well as the plasticizer caprolactam or the stimulant caffeine were determined at high ng/L levels in IWW samples of both WWTPs (see Table 3). Although caprolactam, for example, can be degraded up to 40% in 28 days by the action of certain microorganisms (López Rocha *et al.*, 2020), the adequate elimination of ECs in WWTPs is a crucial issue especially if they are present at such high concentration levels. On the other hand, it has to be mentioned that metformin (recently included in the WL-3) (Gomez Cortes *et al.*, 2020) is by far the most popular diabetes medication worldwide, which has been demonstrated to be hardly metabolized in the human body (Krentz and Bailey, 2005). As a result, it is excreted unaltered and dispersed in wastewater, as has been observed in several studies where the concentration of metformin was non-negligible (Alvarez-Mora *et al.*, 2022; Čelić *et al.*, 2021; Finckh *et al.*, 2022; Golovko *et al.*, 2021). According to the German Umweltbundesamt (UBA) database, such high levels of mycophenolic acid have never been reported, being up to now a concentration of 650 ng/L in surface waters (Franquet-Griell *et al.*, 2017) the highest detected value (https://www.umweltbundesamt.de/en/database-pharmaceuticals-in-the-environment-0, accessed October 2022). The detected large amount of caffeine in untreated samples could be attributed to its high consumption in beverages, as an excipient in a wide variety of drugs and cosmetics. Caffeine concentrations up to 20000 ng/L were reported in the literature (Ebrahimzadeh *et al.*, 2021), but it is eliminated during biological treatment reported (Qi *et al.*, 2015) as it was observed also in this work (> 90% of elimination rate).

Table 2
Target analysis of features identified as level 1 in Crispijana and Galindo WWTPs

Compounds	Abbre-viations	LOQproc (ng/L)	IWW Crispijana WWTP					EWW Crispijana WWTP					IWW Galindo WWTP				
			Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)
2-Hydroxybenzothiazole	OBT	15.4	14	245	845	414	365	15	88	271	146	137	11	1270	2950	1772	1591
4-tert-octylphenol		138.5	5	235	905	470	360	4	159	2238	812	425	0	<LOD	<LOD		
Acetaminophen		2.9	16	7548	24,098	17,275	17,827	3	163	484	293	233	11	31,269	58,474	44,466	47,978
Amantadine		3.3	14	15	31	23	24	15	20	49	35	37	11	45	91	66	62
Amitriptyline		5.4	13	702	2232,000	1081	1019	9	38	1902	482	219	8	22	47	34	33
Atenolol		6	16	173	435	316	311	15	118	236	193	203	11	556	964	770	805
Azithromycin		17.2	0	<LOQproc	<LOQproc			11	25	73	46	43	0	<LOQproc	<LOQproc		
Bendiocarb		6.5	15	9	52	27	24	0	<LOQproc	<LOQproc			9	9	51	22	19
Bentazone		6.2	12	7	9	8	8	13	7	21	14	14	4	20	38	26	23
Benzophenone		0	6	26	53	41	40	0	<LOQproc	<LOQproc			10	39	252	166	165
Bezafibrate		2.9	13	8	16	13	12	16	4	24	14	14	11	181	293	237	252
Bicalutamide		5.4	12	6	15	9	8	16	15	72	45	51	10	6	25	19	20
Bis(2-ethylhexyl) phthalate	DEHP	138.5	6	315	2340	920	733	8	438	1405	1004	984	6	2225	42,815	14,194	7029
Bisoprolol		3.3	7	18	39	25	24	16	21	92	61	59	11	230	327	272	273
Bisphenol A	BPA	15.1	15	362	2709	1719	1727	15	44	400	149	115	10	1098	2702	1717	1696
Bupropion		4.7	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOQproc	<LOQproc		
Caffeine		338.3	16	9587	28,480	20,860	20,858	0	<LOQproc	<LOQproc			11	30,315	82,035	59,811	62,439
Caprolactam		31.9	15	702	2232	1147	1021	8	32	329	152	116	11	18,054	72,388	34,602	29,917
Carbamazepine		6.6	16	20	33	25	24	16	31	176	113	118	11	54	86	68	66
Carbendazim		7.6	16	20	83	52	52	15	15	53	32	29	11	28	104	60	61
Celecoxib		4.2	12	5	11	7	6	15	7	15	10	10	6	10	20	16	18
Cetirizine		4.5	13	5	173	87	90	16	55	252	146	149	4	165	214	196	202
Ciprofloxacin		19.8	13	52	203	118	109	14	29	185	63	56	11	144	327	228	200
Clarithromycin		5.5	0	<LOQproc	<LOQproc			6	40	334	122	83	1	14	14	14	14
Clopidogrel		6.8	3	8	8	8	8	9	8	13	10	10	11	10	19	15	16
Clozapine		3.2	0	<LOQproc	<LOQproc			16	16	99	53	56	2	12	13	13	13
Cotinine		6.2	16	434	1529	971	1023	15	48	264	182	201	11	1626	3288	2381	2381
Dibutyl phthalate	DBP	28.3	13	595	1411	981	971	10	58	286	139	143	11	1262	3263	2093	2041
Diethyl phthalate	DEP	130.6	10	233	1373	674	724	8	326	10,897	4759	2544	11	1819	42,444	6973	3252
Diethyl Toluamide	DEET	6.5	16	24	264	113	75	11	11	86	38	28	11	60	279	165	153
Dioctyl phthalate	DOP	45	6	323	2398	942	751	8	449	1439	1029	1008	6	1186	37,220	12,233	6232
Diuron		5.8	16	34	105	65	68	16	42	206	140	154	11	60	287	109	96
Efavirenz		6.6	6	10	23	15	14	16	18	74	48	51	10	22	63	41	38
Eprosartan		7.4	8	253	834	547	497	14	8	252	73	59	11	1755	3111	2473	2589
Estriol		55.6	9	68	112	90	87	1	72	72	72	72	4	56	145	106	111
Ethyl-S,S-diphenyldithiophosphate	EDDP	3.7	0	<LOD	<LOD			0	<LOD	<LOD			1	8	8	8	8
Finasteride		3.2	0	<LOD	<LOD			0	<LOD	<LOD			10	8	32	18	13
Fluconazole		2.9	15	52	172	108	106	7	36	421	211	167	11	293	1321	658	579
Furosemide		6.5	15	240	623	409	410	14	51	407	232	258	10	186	559	351	331
Gabaspentin		15.3	13	906	3788	2164	2201	15	110	657	453	494	11	2646	5013	4028	3943
Genistein		338.3	13	495	1316	829	777	0	<LOQproc	<LOQproc			11	854	6278	3009	3236
Genistin		6.4	16	64	342	160	149	0	<LOQproc	<LOQproc			9	174	293	222	195

Compounds	EWW1 Galindo WWTP					EWW2 Galindo WWTP					EWW3 Galindo WWTP				
	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)
2-Hydroxybenzothiazole	12	1270	3055	2152	2153	12	135	349	210	211	6	21	142	79	74
4-tert-octylphenol	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOD	<LOD		
Acetaminophen	12	34,750	85,774	63,685	63,832	10	127	380	243	238	9	83	195	140	133
Amantadine	12	71	292	186	200	12	39	68	59	63	12	6	38	15	11
Amitriptyline	7	21	195	81	65	12	34	65	49	51	4	8	19	13	12
Atenolol	12	630	1766	1277	1288.5	12	196	369	303	320	12	126	341	206	195
Azithromycin	0	<LOQproc	<LOQproc			12	390	965	693	719	5	45	547	356	409
Bendiocarb	6	32	99	72	78.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Bentazone	7	10	89	29	19	4	7	14	10	10	1	11	11	11	11
Benzophenone	12	179	1029	439	365	12	33	155	94	96	4	55	103	73	67
Bezafibrate	12	154	361	267	263	12	44	104	76	73	8	6	62	32	27
Bicalutamide	8	19	63	41	39	12	25	53	42	43	12	35	58	47	48
Bis(2-ethylhexyl) phthalate	7	75	4528	1854	644	4	49	12,759	3269	133	2	185	1113	649	649
Bisoprolol	12	235	933	627	680.5	12	225	589	415	400	12	67	527	228	202
Bisphenol A	10	1084	2612	2097	2280	12	134	409	283	298	12	151	341	239	252
Bupropion	4	8	14	11	11	12	6	14	11	11	11	6	18	9	8
Caffeine	12	37,859	136,871	95,781	105,386.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Caprolactam	12	28,020	158,832	89,859	93,068	12	77	2399	619	474	12	189	955	525	488
Carbamazepine	12	86	290	193	214	12	103	204	149	147	5	33	86	56	54
Carbendazim	12	41	222	139	159.5	12	40	82	62	65	4	12	46	22	15
Celecoxib	7	6	15	11	12	12	8	13	10	10	3	6	7	7	7
Cetirizine	4	116	192	157	160	12	120	226	167	158	3	34	57	43	38
Ciprofloxacin	12	20	241	94	71	12	53	116	79	78	3	51	58	55	57
Clarithromycin	2	18	26	22	22	12	16	42	27	28	3	10	17	14	15
Clopidogrel	12	21	63	40	35.5	12	10	19	14	15	0	<LOQproc	<LOQproc		
Clozapine	11	5	116	64	52	12	85	200	132	131	3	10	18	14	14
Cotinine	12	1268	4218	2864	2626	12	164	251	215	225	12	115	218	166	160
Dibutyl phthalate	12	1318	5420	2763	2378	12	108	724	366	371	12	50	403	186	144
Diethyl phthalate	12	1737	37,408	16,053	16,889.5	11	174	7378	2262	903	11	545	11,397	2501	1181
Diethyl Toluamide	12	112	609	322	257.5	12	32	135	77	75	12	29	128	64	53
Dioctyl phthalate	3	849	2064	1626	1965	1	6528	6528	6528	6528	1	148	148	148	148
Diuron	12	62	310	143	139	12	57	97	78	80	12	11	63	35	33
Efavirenz	10	29	57	44	44.5	12	35	57	47	48	9	22	46	34	34
Eprosartan	12	3231	10,449	7012	7235	12	148	470	307	293	4	8	123	64	63
Estriol	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Ethyl-S,S-diphenyldithiophosphate	0	<LOQproc	<LOQproc			12	9	18	14	14	10	6	15	9	8
Finasteride	9	13	46	25	23	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Fluconazole	12	262	1839	934	873.5	12	218	413	337	355	12	199	701	496	557
Furosemide	1	416	416	416	416	11	127	397	266	266	2	19	24	22	22
Gabaspentin	12	5472	16,4408	11,126	11,636.5	12	442	967	670	663	11	76	672	165	132
Genistein	12	1157	14,506	7036	5866.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Genistin	12	117	1216	526	506.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		

Compounds	Abbre-viations	LOQproc (ng/L)														
			Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)
Hydrocortisone		4.4	0	<LOD	<LOD			0	<LOD	<LOD			5	268	631	425	421
Hydroxychloroquine		9.4	0	<LOQproc	<LOQproc			3	32	71	53	57	2	116	132	124	124
Imidacloprid		9.6	3	13	22	17	16	12	12	29	20	20	2	27	47	37	37
Indomethacin		14	9	9	21	14	12	12	17	61	37	35	0	<LOQproc	<LOQproc		
Irbesartan		5.3	6	15	97	64	69	16	80	252	175	183	9	342	582	413	399
Ketoprofen		6	15	115	249	201	205	16	9	103	52	48	11	295	632	525	565
Lidocaine		6	16	15	52	34	35	16	30	166	98	91	11	25	141	94	106
Lopinavir		6.8	4	9	20	14	13	16	7	33	15	12	7	22	121	49	36
Lorazepam		3.2	16	82	203	142	148	7	82	743	354	235	11	264	372	319	324
Losartan		2.9	14	196	550	351	347	7	25	145	87	84	11	449	718	611	639
Mebendazole		3	13	10	32	18	16	15	13	42	29	30	11	35	85	58	59
Mecoprop		5.4	16	56	192	89	82	16	21	489	158	142	0	<LOD	<LOD		
Medroxyprogesterone		3.4	14	41	195	91	82	0	<LOQproc	<LOQproc			6	277	561	388	329
Memantine		9.1	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOQproc	<LOQproc		
Metformin		6.1	14	3930	8780	5953	6193	5	143	287	227	230	11	12,583	21,454	18,525	19,183
Methylparaben		65.7	0	<LOD	<LOD			0	<LOD	<LOD			11	2686	5553	3788	3752
Metoprolol		4.1	1	11	11	11	11	7	11	22	15	14	11	59	392	165	129
Monobutyl phthalate	MBP	16.1	8	268	579	438	472	10	13	592	305	273	1	165	165	165	165
Mycophenolic acid		3.1	15	1211	2701	1762	1670	3	49	55	52	51	11	1819	3457	2993	3103
Naproxen		31.1	0	<LOD	<LOD			0	<LOD	<LOD			11	3739	8712	7461	7969
Nonylphenol		189.5	0	<LOD	<LOD			0	<LOD	<LOD			6	189.5	287	242	252
Norfloxacin		29.6	13	99	350	178	164	13	30	328	139	94	3	2597	10,008	5678	4430
Ofloxacin		1.8	0	<LOD	<LOD			0	<LOD	<LOD			11	80	124	96	92
Omeprazole		3.2	7	10	92	49	27	16	35	123	77	85	7	29	88	56	59
Pentoxifylline		6.3	16	25	65	36	33	8	9	41	18	15	11	111	229	166	161
Perfluorobutanesulfonic acid	PFBS	3	3	7	10	9	9	15	9	602	59	21	10	7	120	23	12
Perfluorooctanoic acid	PFOA	4.5	8	63	109	82	82	14	70	262	167	178	0	<LOD	<LOD		
Pravastatin		6.8	11	146	354	250	251	0	<LOQproc	<LOQproc			0	<LOD	<LOD		
Primidone		5.2	0	<LOQproc	<LOQproc			7	124	422	247	221	0	<LOQproc	<LOQproc		
Propamocarb		17.9	0	<LOD	<LOD			0	<LOD	<LOD			11	93	554	250	221
Propiconazole		5.8	15	7	24	13	11	6	19	89	44	34	4	7	25	13	9
Propyphenazone		2.6	14	5	8	7	7	15	10	26	19	20	11	21	57	36	33
Pyrantel		4	0	<LOD	<LOD			0	<LOD	<LOD			4	22	76	46	44
Ritonavir		32.5	1	38	38	38	38	0	<LOQproc	<LOQproc			7	40	105	72	69
Ropinirole		5.9	1	6	6	6	6	6	9	29	17	14	0	<LOQproc	<LOQproc		
Sertraline		4.4	1	4	4	4	4	14	7	17	10	10	0	<LOQproc	<LOQproc		
Sotalol		5.9	16	9	20	15	15	15	14	24	20	22	11	235	858	465	437
Sulfadiazine		6.1	0	<LOD	<LOD			0	<LOD	<LOD			5	8	32	20	14
Sulfamethoxazole		4.6	0	<LOD	<LOD			0	<LOD	<LOD			11	92	5301	855	183
Sulfapyridine		6.8	4	18	22	21	22	4	18	26	22	22	8	18	44	29	30
Telmisartan		6.1	13	120	376	243	232	16	6	903	643	701	1	1216	1216	1216	1216
Terbutryn		2.9	15	7	40	21	19	8	11	65	31	23	8	63	115	83	84
Testosterone		2.9	15	36	91	49	48	0	<LOQproc	<LOQproc			11	131	468	239	196
Thiabendazole		4.6	0	<LOQproc	<LOQproc			14	6	17	12	12	11	9	34	17	17
Tramadol		16.7	16	142	326	242	233	16	317	1294	819	851	11	726	1805	1404	1436
Triethylphosphate		2.9	4	4	19	10	8	3	7	153	58	14	10	3	233	62	37
Trimethoprim		2.9	14	11	31	23	25	14	16	60	39	41	11	53	1604	314	93
Triphenylphosphate		6.3	16	16	38	28	28	16	8	14	11	11	3	10	18	13	12
Valsartan		16	16	112	1422	1000	1021	15	53	368	163	182	11	5114	9538	6976	6663

Compounds	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)	Times detected	Min Conc. (ng/L)	Max Conc. (ng/L)	Mean (ng/L)	Median (ng/L)
Hydrochlorothiazide	1	231	231	231	231	6	226	231	281	275	7	18	146	79	69
Hydrocortisone	2	186	286	236	236	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Hydroxychloroquine	11	122	372	213	169	12	65	140	118	126	1	82	82	82	82
Imidacloprid	5	22	42	35	38	11	16	32	26	27	11	15	51	29	27
Indomethacin	0	<LOQproc	<LOQproc			11	11	22	14	13	0	<LOQproc	<LOQproc		
Irbesartan	9	226	554	368	329	12	201	381	335	352	11	6	291	109	63
Ketoprofen	12	279	780	579	602.5	12	68	155	115	106	12	8	64	37	31
Lidocaine	12	63	406	263	270	12	54	145	102	103	5	7	60	29	30
Lopinavir	7	12	68	31	25	12	10	68	23	18	12	9	57	19	14
Lorazepam	12	271	978	619	640.5	12	382	892	605	591	8	47	731	260	221
Losartan	12	590	2064	1365	1412	12	198	439	326	320	5	14	353	118	40
Mebendazole	12	68	128	96	99	12	19	29	24	24	4	7	17	13	13
Mecoprop	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOD	<LOD		
Medroxyprogesterone	3	181	309	258	285	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Memantine	10	32	101	72	72	12	56	128	91	91	4	41	83	58	54
Metformin	12	19,439	83,111	51,417	54,757	12	1216	3921	2173	2079	11	13	7338	1185	269
Methylparaben	12	2899	11,642	7375	7203.5	1	443	443	443	443	2	132	151	142	142
Metoprolol	12	74	889	295	256.5	12	39	129	59	56	9	5	51	23	21
Monobutyl phthalate	5	5	650	272	199	12	5	590	150	110	12	31	270	72	49
Mycophenolic acid	12	2591	10,739	7157	7975.5	10	47	170	91	81	4	4	5	4	4
Naproxen	12	4164	9640	7370	7558.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Nonylphenol	9	189.5	440	316	340	10	203	240	219	216	12	200	305	255	256
Norfloxacin	3	1805	15,929	7152	3723	9	44	2373	446	107	12	30	4986	624	86
Ofloxacin	3	64	82	73	74	10	33	74	56	58	0	<LOQproc	<LOQproc		
Omeprazole	12	35	123	72	72	12	29	106	58	51	1	9	9	9	9
Pentoxifylline	12	151	693	465	521	12	46	139	98	111	12	20	198	93	95
Perfluorobutanesulfonic acid	11	3	335	49	21	11	6	49	12	8	12	7	37	11	8
Perfluorooctanoic acid	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOD	<LOD		
Pravastatin	0	<LOD	<LOD			0	<LOD	<LOD			0	<LOD	<LOD		
Primidone	0	<LOQproc	<LOQproc			12	77	319	188	179	12	38	384	232	273
Propamocarb	12	58	534	234	202	2	38	51	45	45	0	<LOQproc	<LOQproc		
Propiconazole	6	11	45	20	17.5	9	7	16	11	12	12	8	28	15	14
Propyphenazone	12	41	115	83	90	12	19	41	32	33	0	<LOQproc	<LOQproc		
Pyrantel	12	35	166	96	102	12	42	86	65	64	6	14	46	27	25
Ritonavir	8	33	102	57	50.5	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Ropinirole	0	<LOQproc	<LOQproc			10	7	21	11	9	7	7	14	10	8
Sertraline	0	<LOQproc	<LOQproc			12	10	25	18	18	0	<LOQproc	<LOQproc		
Sotalol	12	185	681	438	482	12	79	113	99	100	2	56	68	62	62
Sulfadiazine	3	10	36	19	12	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Sulfamethoxazole	12	115	7306	1525	311.5	12	55	1354	308	83	1	62	62	62	62
Sulfapyridine	4	36	51	42	40	1	18	18	18	18	0	<LOQproc	<LOQproc		
Telmisartan	0	<LOQproc	<LOQproc			12	1111	1658	1469	1493	11	26	1088	476	269
Terbutryn	6	90	326	210	235	12	66	126	94	95	3	8	26	15	11
Testosterone	10	111	375	214	205	9	16	692	180	128	0	<LOQproc	<LOQproc		
Thiabendazole	11	14	59	36	37	12	14	42	28	28	8	6	30	17	16
Tramadol	12	1350	6444	4011	4035.5	12	1402	2537	2087	2127	9	33	1020	370	129
Triethylphosphate	12	16	269	99	72	12	54	141	93	94	12	34	110	77	80
Trimethoprim	12	29	1334	281	83	12	37	598	166	68	1	16	16	16	16
Triphenylphosphate	2	7	13	10	10	0	<LOQproc	<LOQproc			0	<LOQproc	<LOQproc		
Valsartan	12	3425	9402	6092	5822	12	102	501	268	243	12	80	746	290	200

Table 3
Qualitative comparison between compounds detected during COVID-19 lockdown and pre-pandemic in the secondary effluent of Galindo WWTP

Class of compound	Compounds detected during COVID-19	Use	Identification level	Detected pre-COVID-19
Drugs used in COVID-19 treatment	Acetaminophen	Pharmaceutical/Analgesic	1	Yes
	Azithromycin	Pharmaceutical/Antibiotic	1	Yes
	Hydroxychloroquine	Pharmaceutical/Antimalarial	1	No
	Lopinavir	Pharmaceutical/Antiretroviral	1	No
	Darunavir	Pharmaceutical/Antiretroviral	2a	Yes
Other related pharmaceuticals	Amantadine	Pharmaceutical/Antiviral	1	Yes
	Amitriptyline	Pharmaceutical/Antidepressant	1	Yes
	Atenolol	Pharmaceutical/Antihypertensive	1	Yes
	Bisoprolol	Pharmaceutical/Antihypertensive	1	Yes
	Candesartan	Pharmaceutical/Antihypertensive	2a	No
	Carbamazepine	Pharmaceutical/Anticonvulsant	1	Yes
	Celiprolol	Pharmaceutical/Antihypertensive	2a	No
	Ciprofloxacin	Pharmaceutical/Antibiotic	1	No
	Citalopram	Pharmaceutical/Antidepressant	3	Yes
	Clarithromycin	Pharmaceutical/Antibiotic	1	No
	Clozapine	Pharmaceutical/Antipsychotic	1	No
	Doxylamine	Pharmaceutical/Anti-inflammatory	2a	Yes
	Efavirenz	Pharmaceutical/Antiretroviral	1	Yes
	Enalaprilat	Pharmaceutical/Antihypertensive	2a	Yes
	Eprosartan	Pharmaceutical/Antihypertensive	1	No
	Fluconazole	Pharmaceutical/Antifungal	1	Yes
	Indomethacin	Pharmaceutical/Anti-inflammatory	1	No
	Irbesartan	Pharmaceutical/Antihypertensive	1	Yes
	Ketoprofen	Pharmaceutical/Anti-inflammatory	1	No
	Lacosamide	Pharmaceutical/Anticonvulsant	2a	Yes
	Lorazepam	Pharmaceutical/Anxiolytic	1	Yes
	Lormetazepam	Pharmaceutical/Anxiolytic	2a	Yes
	Losartan	Pharmaceutical/Antihypertensive	1	Yes
	Metoprolol	Pharmaceutical/Antihypertensive	1	Yes
	Mexedrone	Pharmaceutical/Antidepressant	2a	No
	Minoxidil	Pharmaceutical/Antihypertensive	2a	No
	Mycophenolic acid	Pharmaceutical/Antibiotic	1	Yes

Class of compound	Compounds detected during COVID-19	Use	Identification level	Detected pre-COVID-19
Other related pharmaceuticals *(Cont.)*	Nalbuphine	Pharmaceutical/Analgesic	2a	No
	Norfloxacin	Pharmaceutical/Antibiotic	1	No
	Oxazepam	Pharmaceutical/Anxiolytic	3	Yes
	Ofloxacin	Pharmaceutical/Antibiotic	1	No
	Primidone	Pharmaceutical/Anticonvulsant	1	No
	Propyphenazone	Pharmaceutical/Anti-inflammatory	1	Yes
	Sertraline	Pharmaceutical/Antidepressant	1	Yes
	Sotalol	Pharmaceutical/Antihypertensive	1	Yes
	Sulfamethoxazole	Pharmaceutical/Antibiotic	1	Yes
	Sulpiride	Pharmaceutical/Antidepressant	2a	No
	Telmisartan	Pharmaceutical/Antihypertensive	1	Yes
	Temazepam	Pharmaceutical/Anxiolytic	2a	Yes
	Tiapride	Pharmaceutical/Antipsychotic	2a	No
	Tramadol	Pharmaceutical/Analgesic	1	Yes
	Trazodone	Pharmaceutical/Antidepressant	2a	Yes
	Trimethoprim	Pharmaceutical/Antibiotic	1	Yes
	Valsartan	Pharmaceutical/Antihypertensive	1	Yes
	Venlafaxine	Pharmaceutical/Antidepressant	2a	Yes
Other related compounds	Amphetamine	Illicit drug	3	Yes
	Cocaine	Illicit drug	2a	No
	Cotinine	Nicotine metabolite	1	No
	Ketamine	Illicit drug	2a	Yes
	Metamphetamine	Illicit drug	3	Yes

After the secondary treatments a removal rate higher than 50% was determined for 22 and 30 compounds (in Crispijana and Galindo WWTP, respectively), and the efficiency of the tertiary treatment from Galindo WWTP was evidenced. By the use of the tertiary treatment, a large number of compounds (n = 32) were significantly removed (see Table S4 in SI). A non-significant elimination rate was observed through the secondary treatment for the rest of identified compounds (i.e., 45 compounds), so that they can be cathegorized as "pseudo-persistent" contaminants that are continuously released into the aquatic ecosystem (see Table S4 in SI).

3.2. *Influence of the COVID-19*

The lack of knowledge of the virus and the need to rapidly find some effective treatments to combat the virus led to the massive use of several pharmaceutical compounds (or combinations) with antiviral and/or antimicrobial activity (Costanzo *et al.*, 2020). In this work, suspect analysis enabled the identification (at level 1 and 2a) of some of those drugs that were massively used for COVID-19 treatment early in the pandemic thereby increasing their occurrence in wastewaters (see Table 4) (Alygizakis *et al.*, 2021; Cappelli *et al.*, 2022;

Galani *et al.*, 2021). Based on some previous occurrence data get in sampling campaigns before COVID-19 time in secondary effluent of Galindo WWTP (González-Gaya *et al.*, 2021), the analgesic acetaminophen, the antibiotic azithromycin, the antivirals darunavir and lopinavir, and the antimalarial hydroxychloroquine are some of those drugs with significant occurrence during the pandemic time.

As can be observed in Table 4, there is no prior evidence of the occurrence of the compounds hydroxychloroquine and lopinavir above detection limits, being the first time that the presence of hydroxychloroquine was registered in Basque environmental waters (Domingo-Echaburu *et al.*, 2022). Hydroxychloroquine, typically used for malaria, lupus and rheumatoid arthritis treatment (Drug Bank Online, 2020), was considered as a possible efficient drug to treat COVID-19 disease (either alone or in combination with azithromycin) at the beginning of the pandemic (Gautret *et al.*, 2020). The use of lopinavir (an antiviral often prescribed with ritonavir to treat HIV (Osborne *et al.*, 2020) as an effective virus-fighting agent was also revealed by its high occurrence in wastewaters during the pandemic period. In fact, according to the UBA, the concentration found for lopinavir in the analyzed samples was the highest registered at the European level (https://www.umweltbundesamt.de/en/database-pharmaceuticals-in-the-environment-0, accessed October 2022). Acetaminophen, typically used in WBE to predict disease outbreaks because it is a short-term application analgesic that can be consumed without prescription (Halwatura *et al.*, 2022), was also used to control some of the COVID-19 symptoms, and hence, its occurrence was detected during the pandemic time but also before that period (see Table 4) (González-Gaya *et al.*, 2021). A similar trend was also observed for the previously highlighted azithromycin and darunavir compounds, which were detected during and before pandemic time (González-Gaya *et al.*, 2021).

Regarding the antibiotics detected in samples collected in this study, although their occurrence is positively correlated with the COVID-19 metrics and it is known that they were massively administered during lockdown (Cappelli *et al.*, 2022; Galani *et al.*, 2021; Gonzalez-Zorn, 2021), the presence of broad-spectrum class antibiotics in wastewaters could be a consequence of seasonal diseases. Heterogeneous trend in pharmaceuticals for other therapeutic purposes (e.g. antihypertensives, anti-inflammatories, anticonvulsants) consumption during the pandemic has been reported. On the other hand, post-traumatic stress, depression, insomnia, fear and/or frustration, among others suffered by citizens during the lockdown (Brooks *et al.*, 2020) (Singh *et al.*, 2020) could led to the consumption of illicit drugs. Qualitative comparison of compounds' occurrence before (González-Gaya *et al.*, 2021) and during the pandemic time (this study) revealed negligible differences in the presence of most of the compounds detected in this study at the Galindo WWTP, with only 20 (e.g. hydroxychloroquine, lopinavir, clarithromycin, clozapine, sulpiride and tiapride, among others) compounds more detected in samples collected during the lockdown (see Table 4); particularly, new pharmaceuticals have emerged in Galindo WWTP effluent (e.g., candesartan, clozapine, eprosartan or primidone, among others). In line with other studies (Alygizakis *et al.*, 2021; Nason *et al.*, 2022; Wang *et al.*, 2020), a higher number of antipsychotic drugs (including antidepressants) have been observed compared to the non-COVID-19 period, which, as aforementioned, would give more insight into the mental health of the Basque citizens provoked by the different measures applied. Furthermore, certain illicit drugs considered as biomarkers in WBE studies (Alygizakis *et al.*, 2021; Been *et al.*, 2021; Reinstadler *et al.*, 2021) such as amphetamine or ketamine were also detected (see Table 4).

Unfortunately, the lack of previous studies hindered the comparison of the values detected at the Crispijana WWTP. However, an increase in hospital drug consumption of certain selected drugs during the first wave pandemic was previously discussed (Domingo-Echaburu *et al.*, 2022).

3.3. *Prioritization strategy for environmental risk assessment*

A prioritization strategy for environmental risk assessment was carried out using the compounds quantified in the effluents of Crispijana and Galindo WWTPs. The compounds were scored based on the (a) removal efficiency (RE, %), (b) estimated persistency (half-life time in days, DT50), (c) bioconcentration factor (BCF), (d) toxicity potential and (e) frequency of detection in the samples (see section 2.8). Those compounds with the lowest total score value were set as the potential drivers of toxicity.

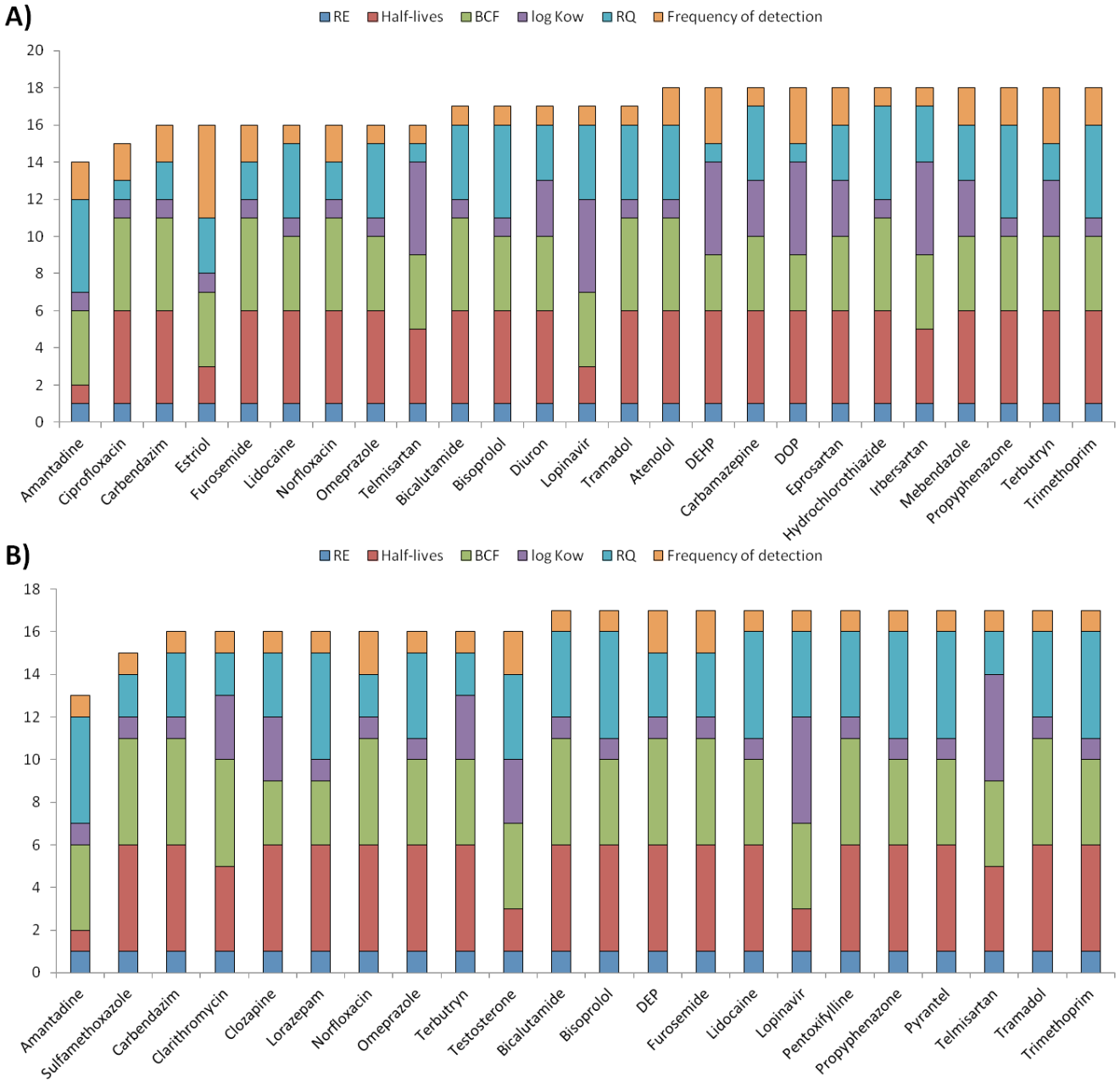


Figure 2. Total scores of the top risk drivers found in the secondary effluent of Crispijana (A) and Galindo WWTPs (B)

Among the compounds quantified in both WWTPs, the list of the most concerning compounds is constituted by 25 and 22 micropollutants in Crispijana and Galindo, respectively. Pharmaceutical compounds dominated both priority lists (> 70% of the total in both WWTPs), while, lower total scores were obtained in wastewaters from Galindo WWTP for the prioritized contaminants (total score ≤ 17 vs 18) (see Figure 2, Table S6 in SI). Several compounds identified as priority compounds in this work have already been considered hazardous elsewhere such as the ones included in WFD priority list (DEHP, diuron and terbutryn) (European Commission, 2013) and the ones included in the current Watch List to be considered for future prioritization (clarithromycin and sulfamethoxazole) (European Commission, 2015; Gomez Cortes *et al.*, 2020). Moreover, some of the compounds considered in here as priority compounds were also pointed out as key chemicals in environmental toxicity studies. In the work of Gros and coworkers, for example, lidocaine (included in both priority rankings) was pointed out as one of the top-risk drivers of Swedish wastewaters, followed by diuron (included in the priority list of Crispijana WWTP) to a lower extent (higher total scores) (Gros *et al.*, 2017). Carbamazepine, irbesartan, sulfamethoxazole and ciprofloxacin were identified as relevant chemicals for marine organisms in the area of Ebro Delta (Spain) in the work of Čelić and coworkers, where a similar prioritization strategy to the one used in the present work was done (Čelić *et al.*, 2019). After the assessment of 52 European WWTPs, Finckh *et al.* pointed out carbendazim, terbutryn and diuron as toxicity-driver compounds (Finckh *et al.*, 2022). Moreover, other recent studies based on the calculation of RQs in WWTP effluents (Figuière *et al.*, 2022; Lopez-Herguedas *et al.*, 2022; Solaun *et al.*, 2021), freshwater (Figuière *et al.*, 2022) and riverine and coastal ecosystems (Čelić *et al.*, 2021) highlighted the need to prioritize some of the concerning compounds pointed out in the present work.

Secondary treatments implemented in both analyzed WWTPs seemed to be not efficient enough to remove completely all the prioritized contaminants (score of 1). The poor elimination rate of the detected organic micropollutants through conventional secondary treatments implemented in WWTPs is widely reported in the literature (Golovko *et al.*, 2021; Jelic *et al.*, 2011; Köck-Schulmeyer *et al.*, 2013; Kovalova *et al.*, 2012; Le Corre *et al.*, 2012). The associated matrix effect that can result in signal suppression is usually the argument used to explain these "negative" removals. However, typical retransformation of conjugated compounds into the original compound through biological processes, improper sample collection (lack of correlation between influent and effluent samples due to a bad timely collection) or the release of the compounds from fecal particles due to microbial breakdown can also be considered to report negative compound removals (Fernández-López *et al.*, 2016; Köck-Schulmeyer *et al.*, 2013).

Amantadine (score 1) and lopinavir (score 2) stood out as the most persistent compounds in both WWTPs, showing DT50 values exceeding 60 days, with the addition of estriol (Crispijana WWTP, score 2) and testosterone (Galindo WWTP, score 2). The persistency of the remaining compounds was lower (< 37.5 days), suggesting that most of the top compounds were easily degradable (see Figure 2, Table S6 in SI). DEHP and DOP in Crispijana WWTP and clozapine and lorazepam in Galindo WWTP were the compounds showing the highest predicted BCF values, however, none of the detected compounds could be considered as highly bioaccumulative (BCF < 100). Additionally, it is important to note that statements made considering biodegradation and bioaccumulation of the compounds are fully based on predicted values due to the lack of experimental values and contradictions may exist, as was observed when comparing half-life times and REs. Thus, there could be an overestimation of the real risk. In consequence, these categories should not share the same weight as categories based on experimental data in future prioritization strategies.

In terms of mobility, prioritized compounds showed, overall, low log K_{ow} values, suggesting a high mobility potential, with the exception of DOP, irbesartan, lopinavir and telmisartan (see **F**igure 2, Table S6 in SI).

Individual RQs were calculated to assess the maximum concentration at which the ecological status of the ecosystem is preserved. To that aim, predicted values based on in-silico tools (i.e. ECOSAR) for baseline toxicity were considered, since there is a lack of experimental toxicity data available for the assessed compounds (see Table S5). In this case, experimental toxicity values were found for around 50 and 60% of the prioritized compounds for PNEC calculation in Crispijana and Galindo WWTPs, respectively. Estimated individual toxicities highlighted that although most of the detected compounds do not pose a relevant environmental risk, some compounds should be closely tracked, especially ciprofloxacin, telmisartan, DEHP and DOP (RQ > 1), and sulfamethoxazole, clarithromycin, norfloxacin and terbutryn (RQ > 0.1), in a lesser extent. Furthermore, the over/underestimation of the environmental risk led by the use of predicted ecotoxicological data rather than experimental (i.e. NOEC and/or EC_{50}) for the calculation of RQs emphasizes the need for more empirical evidence to provide more reliable results.

Both priority rankings include compounds that have not been identified in previous studies as concerning and which may be related in some way to COVID-19 disease. Lopinavir, as aforementioned, has been used in combination with ritonavir to combat the virus, suggesting that its massive use during this particular period is responsible for increasing the potential environmental risk it may pose. On the other hand, the potential risk of the psychoactive compounds clozapine and lorazepam could be correlated with their raised prescription rates to overcome mental illnesses caused by the lockdown.

Comparing both secondary effluents with the tertiary effluent of Galindo WWTP, slightly higher total scores of the top-ranked contaminants were obtained in the latter (see section S7 in SI).

Table 4
Potential antimicrobial and antiviral activity of the drugs of interest in both analyzed WWTPs

Compounds	PNEC-AR (µg/L) (Bengtsson-Palme and Larsson, 2016	vIC50/vEC50 (µg/L) (Kuroda *et al.*, 2021)	Crispijana WWTP		Galindo WWTP	
			RQ-AR	EDRP	RQ-AR	EDRP
Ciprofloxacin	0.064		0.1742	—	0.0347	—
Clarithromycin	0.25		0.0676	—	0.00314	—
Fluconazole	0.25		0.1411	—	0.032644	—
Hydroxychloroquine		242	—	0.000025	—	1.14339E-05
Lopinavir		1088	—	2.96415E-06	—	9.26471E-07
Norfloxacin	0.5		0.05105	—	0.065858	—
Ofloxacin	0.5		—	—	0.002938	—
Ritonavir		6222	—	—	—	1.04468E-07
Sulfamethoxazole	16,000		—	—	0.001536438	—
Trimethoprim	0.5		0.0105	—	0.023326	—

Considering the high loads of pharmaceuticals with antimicrobial and antiviral activity released into the environment due to the COVID-19 disease, the concern of the development of resistance in the aquatic environment has increased (Knight *et al.*, 2021; Kuroda *et al.*, 2021). The antimicrobial and antiviral potential activity of the drugs of interest was determined with the calculation of RQ-AR and EDRP (see section 2.8). The risk indices determined (see Table 5) suggest that none of the detected compounds might pose a relevant activity, since RQ-AR and EDRP values did not exceed the threshold of

> 1. However, in the case of antimicrobial activity, ciprofloxacin and fluconazole reached concentrations of medium antimicrobial resistance risk (1 > RQ-AR > 0.1). Our findings, considering the antimicrobial activity, were contrary to those observed by Cappelli and coworkers, as in that case both azithromycin and ciprofloxacin exceeded the RQ-AR = 1 threshold, posing a high potential for developing antimicrobial resistance (Cappelli *et al.*, 2022). Nevertheless, it should be highlighted that any DF (see section 2.8) was applied in that study, representing the worst-case scenario. On the other hand, the negligible risk of EDRP determined in this study is in line with other studies (Cappelli *et al.*, 2022; Kuroda *et al.*, 2021). However, regardless of the determined low RQ-AR and EDRP values, a reduction of antiviral and antimicrobial drug residues is suggested in order to avoid the disruption of natural biological systems as well as the development of resistance in aquatic systems (Kuroda *et al.*, 2021; Usman *et al.*, 2020).

Once the priority list of contaminants was defined, mixture toxicity was assessed via the calculation of STU (see section 2.8). All effluent samples exceeded the threshold of 1 (Figure 3) obtaining the highest mixture risk (STU = 11.1) for the secondary effluent of Crispijana WWTP being DOP the main contributor of the mixture toxicity (72% of the total) followed by DEHP and telmisartan (STU values of 1.28 and 1.11, respectively). In the case of the secondary effluent of Galindo WWTP, the risk was almost halved to an STU value of 6.8, predominated by DEHP which contributed to around 90% of the total mixture risk, while more than the remaining mixture toxicity was attributed to norfloxacin. Similarly to the individual risk assessment, the lowest STU value was estimated for the tertiary effluent of Galindo WWTP (STU = 1.6). In this latter case, any of the compounds exceeded the threshold of 1 being DEHP and norfloxacin the most influential compounds in the mixture risk both with moderate risks (0.63 and 0.79, respectively).

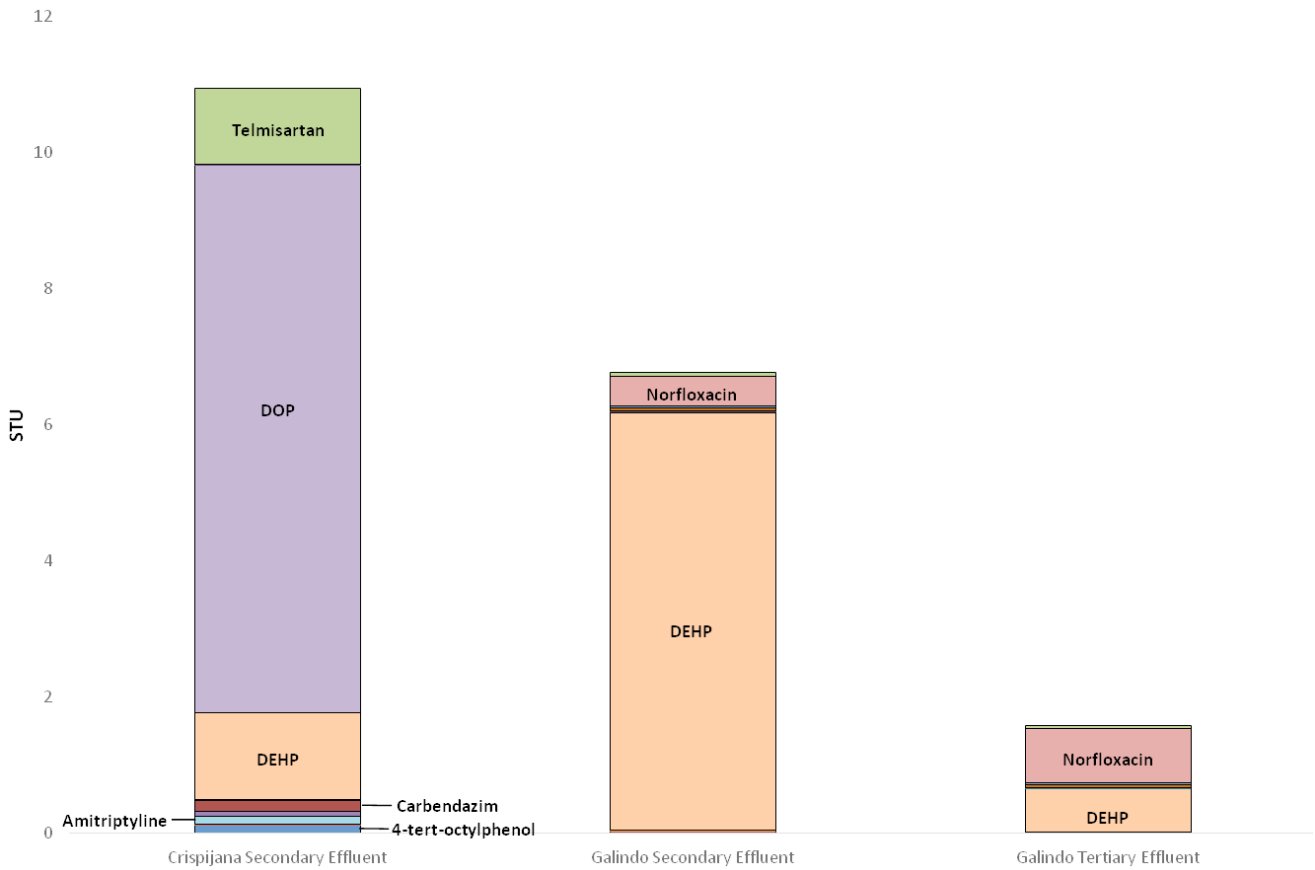


Figure 3. STU values for analyzed effluent samples including the main contributors

Chronic ecotoxicological data was considered rather than acute data when possible for the mixture toxicity assessment (see section 2.8). As indicated by Markert *et al.* the choice of acute or chronic toxicity data will have a clear impact on the calculated risks of the mixture, and they recommend that the risk assessment of the mixture should be based not only on the commonly applied acute toxicity data but also on the chronic toxicity data (Markert *et al.*, 2020). In fact, with many of the contaminants, it is known that it is the long-term risks that will really affect the environment. However, the use of fixed ratios for the extrapolation from acute to chronic toxicity is problematic, because some chemicals show different modes of action (MoA) under short- and long-term conditions (Ahlers *et al.*, 2006). In addition, the biological mechanisms of action differ from species to species.

4. Conclusions

A previously validated suspect screening workflow was used for the identification of emerging contaminants present in two different WWTPs located in the Basque Country (Crispijana and Galindo) during COVID-19 confinement. Pharmaceutical compounds used for COVID-19 disease treatment were detected in both WWTP samples including the antivirals ritonavir/lopinavir (level 1) and darunavir (level 2a), the antimalarial hydroxychloroquine (level 1) and the antibiotic azithromycin (level 1). Moreover, other pharmaceuticals used for therapeutic purposes were also detected (e.g. amitriptyline, clozapine, lorazepam, primidone and valsartan, among others), suggesting a positive correlation with the mental illnesses caused by the lockdown. Despite the differences between the number and concentrations of the compounds found in both WWTPs due to their different locations, the population of influence and the treatments implemented, they both coincide in not being able to eliminate most of the drugs found in their influents with any of the treatments implemented.

A prioritization strategy for the ECs detected in WWTP effluent samples was carried out in order to point out the major contributors to environmental risk. Although several compounds were considered of concern, both prioritization lists consisted mostly of pharmaceutical compounds (e.g. amantadine, telmisartan, lopinavir, clarithromycin, clozapine) highlighting the need for monitoring and thereby concluding whether they should be considered for future regulation. On the other hand, the lack of measured data (e.g. degradation, bioaccumulation and toxicity) for many frequently detected compounds leaves no alternative but to make use of reference QSARs or other in-silico tools for data prediction, which leads to high uncertainty in the affirmations made. Although the values determined to assess antimicrobial and antiviral resistance activity for the compounds of interest were low (RQ-AR and EDRP values < 1), the results of the antimicrobial risk index showed medium environmental concern for the detected levels of ciprofloxacin and fluconazole, demonstrating the need to include these endpoints in current regulatory systems.

Thus, the development of new technologies in the wastewater treatments is required to improve the removal efficiency of those compounds so the potential environmental risk they may pose in receiving water ecosystems decreases. On the other hand, more efforts need to be made to fill the gaps by prioritizing chemicals for effect testing and evaluating the mixture effects (i.e. synergic or antagonistic effects) of the contaminants.

4.1. Acknowledgements

This study was funded by the Basque Government through financial support as a consolidated group of the Basque Research System (IT1446-22), the Agencia Estatal de Investigación (AEI) of Spain, the 2020 call for the generation of knowledge and scientific and technological strengthening of the R&D&i system and the R&D&i focused on society's challenges, through project PID2020-117686RB-C31 and the Council of Vitoria-Gasteiz and Fundación Vital. The authors are grateful to the Consorcio de Aguas de Bilbao and especially to Iñigo González. Naroa Lopez-Herguedas is grateful to the Spanish Ministry of Economy, Industry and Competitivity for her pre-

doctoral scholarship FPI 2018. Iker Alvarez-Mora is grateful to the University of the Basque Country and the Université de Pau et des Pays de L' Adour for his cotutelle predoctoral scholarship. Finally, the authors acknowledge support from the AEI and the Ministry of Science, Innovation and Universities (MICIU) to support the Thematic Network of Excellence (NET-4SEA) on emerging contaminants in marine settings (CTM2017-90890-REDT, MICIU/AEI/FEDER, EU).

5. References

Afonso-Olivares, C., Čadková, T., Sosa-Ferrera, Z., Santana-Rodríguez, J.J., Nováková, L., 2017. Simplified solid-phase extraction procedure combined with liquid chromatography tandem–mass spectrometry for multiresidue assessment of pharmaceutical compounds in environmental liquid samples. J. Chromatogr. A 1487, 54-63. https://doi.org/10.1016/j.chroma.2017.01.059

Ahlers, J., Riedhammer, C., Vogliano, M., Ebert, R.-U., Kühne, R., Schüürmann, G., 2006. Acute to chronic ratios in aquatic toxicity-variation across trophic levels and relationship with chemical structure. Environ. Toxicol. Chem. 25, 2937-2945. https://doi.org/10.1897/05-701R.1

Alvarez-Mora, I., Bolliet, V., Lopez-Herguedas, N., Castro, L., Anakabe, E., Monperrus, M., Etxebarria, N., 2022. Prioritization based on risk assessment to study the bioconcentration and biotransformation of pharmaceuticals in glass eels (Anguilla anguilla) from the Adour estuary (Basque Country, France). Environ. Pollut. 120016. https://doi.org/10.1016/j.envpol.2022.120016

Alygizakis, N., Galani, A., Rousis, N.I., Aalizadeh, R., Dimopoulos, M.-A., Thomaidis, N.S., 2021. Change in the chemical content of untreated wastewater of Athens, Greece under COVID-19 pandemic. Sci. Total Environ. 799, 149230. https://doi.org/10.1016/j.scitotenv.2021.149230

Andersson, D.I., Hughes, D., 2012. Evolution of antibiotic resistance at non-lethal drug concentrations. Drug Resist. Updat. 15, 162-172. https://doi.org/10.1016/j.drup.2012.03.005

Backhaus, T., Faust, M., 2012. Predictive Environmental Risk Assessment of Chemical Mixtures: A Conceptual Framework. Environ. Sci. Technol. 46, 2564-2573. https://doi.org/10.1021/es2034125

Bandala, E.R., Kruger, B.R., Cesarino, I., Leao, A.L., Wijesiri, B., Goonetilleke, A., 2021. Impacts of COVID-19 pandemic on the wastewater pathway into surface water: A review. Sci. Total Environ. 774, 145586. https://doi.org/10.1016/j.scitotenv.2021.145586

Been, F., Emke, E., Matias, J., Baz-Lomba, J.A., Boogaerts, T., Castiglioni, S., Campos-Mañas, M., Celma, A., Covaci, A., de Voogt, P., Hernández, F., Kasprzyk-Hordern, B., Laak, T. ter, Reid, M., Salgueiro-González, N., Steenbeek, R., van Nuijs, A.L.N., Zuccato, E., Bijlsma, L., 2021. Changes in drug use in European cities during early COVID-19 lockdowns - A snapshot from wastewater analysis. Environ. Int. 153, 106540. https://doi.org/10.1016/j.envint.2021.106540

Bengtsson-Palme, J., Larsson, D.G.J., 2016. Concentrations of antibiotics predicted to select for resistant bacteria: Proposed limits for environmental regulation. Environ. Int. 86, 140-149. https://doi.org/10.1016/j.envint.2015.10.015

Bijlsma, L., Picó, Y., Andreu, V., Celma, A., Estévez-Danta, A., González-Mariño, I., Hernández, F., López de Alda, M., López-García, E., Marcé, R.M., Miró, M., Montes, R., Pérez de San Román-Landa, U., Pitarch, E., Pocurull, E., Postigo, C., Prieto, A., Rico, A., Rodil, R., Valcárcel, Y., Ventura, M., Quintana, J.B., 2021. The embodiment of wastewater data for the estimation of illicit drug consumption in Spain. Sci. Total Environ. 772, 144794. https://doi.org/10.1016/j.scitotenv.2020.144794

Booth, A., Aga, D.S., Wester, A.L., 2020. Retrospective analysis of the global antibiotic residues that exceed the predicted no effect concentration for antimicrobial resistance in various environmental matrices. Environ. Int. 141, 105796. https://doi.org/10.1016/j.envint.2020.105796

Boxall, A.B.A., Rudd, M.A., Brooks, B.W., Caldwell, D.J., Choi, K., Hickmann, S., Innes, E., Ostapyk, K., Staveley, J.P., Verslycke, T., Ankley, G.T., Beazley, K.F., Belanger, S.E., Berninger, J.P., Carriquiriborde, P., Coors, A., DeLeo, P.C., Dyer, S.D., Ericson, J.F., Gagné, F., Giesy, J.P., Gouin, T., Hallstrom, L., Karlsson, M.V., Larsson, D.G.J., Lazorchak, J.M., Mastrocco, F., McLaughlin, A., McMaster, M.E., Meyerhoff, R.D., Moore, R., Parrott, J.L., Snape, J.R., Murray-Smith, R., Servos, M.R., Sibley, P.K., Straub, J.O., Szabo, N.D., Topp, E., Tetreault, G.R., Trudeau, V.L., Van Der Kraak, G., 2012. Pharmaceuticals and Personal Care Products in the Environment: What Are the Big Questions? Environ. Health Perspect. 120, 1221-1229. https://doi.org/10.1289/ehp.1104477

Brooks, S.K., Webster, R.K., Smith, L.E., Woodland, L., Wessely, S., Greenberg, N., Rubin, G.J., 2020. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet 395, 912-920. https://doi.org/10.1016/S0140-6736(20)30460-8

Cappelli, F., Longoni, O., Rigato, J., Rusconi, M., Sala, A., Fochi, I., Palumbo, M.T., Polesello, S., Roscioli, C., Salerno, F., Stefani, F., Bettinetti, R., Valsecchi, S., 2022. Suspect screening of wastewaters to trace anti-COVID-19 drugs: Potential adverse effects on aquatic environment. Sci. Total Environ. 824, 153756. https://doi.org/10.1016/j.scitotenv.2022.153756

Čelić, M., Gros, M., Farré, M., Barceló, D., Petrović, M., 2019. Pharmaceuticals as chemical markers of wastewater contamination in the vulnerable area of the Ebro Delta (Spain). Sci. Total Environ. 652, 952-963. https://doi.org/10.1016/j.scitotenv.2018.10.290

Čelić, M., Jaén-Gil, A., Briceño-Guevara, S., Rodríguez-Mozaz, S., Gros, M., Petrović, M., 2021. Extended suspect screening to identify contaminants of emerging concern in riverine and coastal ecosystems and assessment of environmental risks. J. Hazard. Mater. 404, 124102. https://doi.org/10.1016/j.jhazmat.2020.124102

Costanzo, M., De Giglio, M.A.R., Roviello, G.N., 2020. SARS-CoV-2: Recent Reports on Antiviral Therapies Based on Lopinavir/Ritonavir, Darunavir/Umifenovir, Hydroxychloroquine, Remdesi-

vir, Favipiravir and other Drugs for the Treatment of the New Coronavirus [WWW Document]. https://doi.org/10.2174/0929867327666200416131117

de Araújo, J.C., Madeira, C.L., Bressani, T., Leal, C., Leroy, D., Machado, E.C., Fernandes, L.A., Espinosa, M.F., Freitas, G.T.O., Leão, T., Mota, V.T., Pereira, A.D., Perdigão, C., Tröger, F., Ayrimoraes, S., de Melo, M.C., Laguardia, F., Reis, M.T.P., Mota, C., Chernicharo, C.A.L., 2022. Quantification of SARS-CoV-2 in wastewater samples from hospitals treating COVID-19 patients during the first wave of the pandemic in Brazil. Sci. Total Environ. 160498. https://doi.org/10.1016/j.scitotenv.2022.160498

Díaz-Garduño, B., Pintado-Herrera, M.G., Biel- Maeso, M., Rueda-Márquez, J.J., Lara-Martín, P.A., Perales, J.A., Manzano, M.A., Garrido-Pérez, C., Martín-Díaz, M.L., 2017. Environmental risk assessment of effluents as a whole emerging contaminant: Efficiency of alternative tertiary treatments for wastewater depuration. Water Res. 119, 136-149. https://doi.org/10.1016/j.watres.2017.04.021

Dimitrov, S.D., Dermen, I.A., Dimitrova, N.H., Vasilev, K.G., Schultz, T.W., Mekenyan, O.G., 2019. Mechanistic relationship between biodegradation and bioaccumulation. Practical outcomes. Regul. Toxicol. Pharmacol. 107, 104411. https://doi.org/10.1016/j.yrtph.2019.104411

Domingo-Echaburu, S., Irazola, M., Prieto, A., Rocano, B., Lopez de Torre-Querejazu, A., Quintana, A., Orive, G., Lertxundi, U., 2022. Drugs used during the COVID-19 first wave in Vitoria-Gasteiz (Spain) and their presence in the environment. Sci. Total Environ. 820, 153122. https://doi.org/10.1016/j.scitotenv.2022.153122

Drug Bank Online, 2020. . Hydroxychloroquine. URL https://go.drugbank.com/drugs/DB01611

Ebrahimzadeh, G., Nodehi, R.N., Alimohammadi, M., Rezaei Kahkah, M.R., Mahvi, A.H., 2021. Monitoring of caffeine concentration in infused tea, human urine, domestic wastewater and different water resources in southeast of Iran- caffeine an alternative indicator for contamination of human origin. J. Environ. Manage. 283, 111971. https://doi.org/10.1016/j.jenvman.2021.111971

Estévez-Danta, A., Bijlsma, L., Capela, R., Cela, R., Celma, A., Hernández, F., Lertxundi, U., Matias, J., Montes, R., Orive, G., Prieto, A., Santos, M.M., Rodil, R., Quintana, J.B., 2022. Use of illicit drugs, alcohol and tobacco in Spain and Portugal during the COVID-19 crisis in 2020 as measured by wastewater-based epidemiology. Sci. Total Environ. 836, 155697. https://doi.org/10.1016/j.scitotenv.2022.155697

European Commission, 2015. Commission Implementing Decision (EU) 2015/495 of 20 march 2015. Establishing a watch list of substances for union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council. Off. J. Eur. Union 78, 20-30.

European Commission, 2013. Directive 2013/39/EU of the European Parliament and the Council of 12. August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy.

European Parliament, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (Text with EEA relevance) Text with EEA relevance.

Fernández-López, C., Guillén-Navarro, J.M., Padilla, J.J., Parsons, J.R., 2016. Comparison of the removal efficiencies of selected pharmaceuticals in wastewater treatment plants in the region of Murcia, Spain. Ecol. Eng. 95, 811-816. https://doi.org/10.1016/j.ecoleng.2016.06.093

Figuière, R., Waara, S., Ahrens, L., Golovko, O., 2022. Risk-based screening for prioritisation of organic micropollutants in Swedish freshwater. J. Hazard. Mater. 429, 128302. https://doi.org/10.1016/j.jhazmat.2022.128302

Finckh, S., Beckers, L.-M., Busch, W., Carmona, E., Dulio, V., Kramer, L., Krauss, M., Posthuma, L., Schulze, T., Slootweg, J., Von der Ohe, P.C., Brack, W., 2022. A risk based assessment approach for chemical mixtures from wastewater treatment plant effluents. Environ. Int. 164, 107234. https://doi.org/10.1016/j.envint.2022.107234

Fonseca, E., Hernández, F., Ibáñez, M., Rico, A., Pitarch, E., Bijlsma, L., 2020. Occurrence and ecological risks of pharmaceuticals in a Mediterranean river in Eastern Spain. Environ. Int. 144, 106004. https://doi.org/10.1016/j.envint.2020.106004

Franquet-Griell, H., Cornadó, D., Caixach, J., Ventura, F., Lacorte, S., 2017. Determination of cytostatic drugs in Besòs River (NE Spain) and comparison with predicted environmental concentrations. Environ. Sci. Pollut. Res. 24, 6492-6503. https://doi.org/10.1007/s11356-016-8337-y

Gago-Ferrero, P., Schymanski, E.L., Hollender, J., Thomaidis, N.S., 2016. Nontarget Analysis of Environmental Samples Based on Liquid Chromatography Coupled to High Resolution Mass Spectrometry (LC-HRMS). Appl. Time-Flight Orbitrap Mass Spectrom. Environ. Food Doping Forensic Anal. 71, 381-403. https://doi.org/10.1016/bs.coac.2016.01.012

Galani, A., Alygizakis, N., Aalizadeh, R., Kastritis, E., Dimopoulos, M.-A., Thomaidis, N.S., 2021. Patterns of pharmaceuticals use during the first wave of COVID-19 pandemic in Athens, Greece as revealed by wastewater-based epidemiology. Sci. Total Environ. 798, 149014. https://doi.org/10.1016/j.scitotenv.2021.149014

Gautret, P., Lagier, J.-C., Parola, P., Hoang, V.T., Meddeb, L., Sevestre, J., Mailhe, M., Doudier, B., Aubry, C., Amrane, S., Seng, P., Hocquart, M., Eldin, C., Finance, J., Vieira, V.E., Tissot-Dupont, H.T., Honoré, S., Stein, A., Million, M., Colson, P., La Scola, B., Veit, V., Jacquier, A., Deharo, J.-C., Drancourt, M., Fournier, P.E., Rolain, J.-M., Brouqui, P., Raoult, D., 2020. Clinical and microbiological effect of a combination of hydroxychloroquine and azithromycin in 80 COVID-19 patients with at least a six-day follow up: A pilot observational study. Travel Med. Infect. Dis. 34, 101663. https://doi.org/10.1016/j.tmaid.2020.101663

Godini, H., Hoseinzadeh, E., Hossini, H., 2021. Water and wastewater as potential sources of SARS-CoV-2 transmission: a systematic review. Rev. Environ. Health 36, 309-317. https://doi.org/10.1515/reveh-2020-0148

Golovko, O., Örn, S., Sörengård, M., Frieberg, K., Nassazzi, W., Lai, F.Y., Ahrens, L., 2021. Occurrence and removal of chemicals of emerging concern in wastewater treatment plants and their impact on receiving water systems. Sci. Total Environ. 754, 142122. https://doi.org/10.1016/j.scitotenv.2020.142122

Gomez Cortes, L., Marinov, D., Sanseverino, I., Navarro Cuenca, A., Niegowska, M., Porcel Rodriguez, E., Lettieri, T., 2020. Selection of substances for the 3rd Watch List under the Water Framework Directive, EUR 30297 EN. Publications Office of the European Union, Luxembourg.

González-Gaya, B., Lopez-Herguedas, N., Santamaria, A., Mijangos, F., Etxebarria, N., Olivares, M., Prieto, A., Zuloaga, O., 2021. Suspect screening workflow comparison for the analysis of organic xenobiotics in environmental water samples. Chemosphere 274, 129964. https://doi.org/10.1016/j.chemosphere.2021.129964

Gros, M., Blum, K.M., Jernstedt, H., Renman, G., Rodríguez-Mozaz, S., Haglund, P., Andersson, P.L., Wiberg, K., Ahrens, L., 2017. Screening and prioritization of micropollutants in wastewaters from on-site sewage treatment facilities. J. Hazard. Mater. 328, 37-45. https://doi.org/10.1016/j.jhazmat.2016.12.055

Gullberg, E., Albrecht, L.M., Karlsson, C., Sandegren, L., Andersson, D.I., 2014. Selection of a Multidrug Resistance Plasmid by Sublethal Levels of Antibiotics and Heavy Metals. mBio 5, e01918-14. https://doi.org/10.1128/mBio.01918-14

Halwatura, L.M., Mclerran, I.S., Weglarski, D.L., Ahmed, Z.U., Ye, Y., Bradley, I.M., Aga, D.S., 2022. Complementing RNA Detection with Pharmaceutical Monitoring for Early Warning of Viral Outbreaks through Wastewater-Based Epidemiology. Environ. Sci. Technol. Lett. 9, 567-574. https://doi.org/10.1021/acs.estlett.2c00259

Jelic, A., Gros, M., Ginebreda, A., Cespedes-Sánchez, R., Ventura, F., Petrovic, M., Barcelo, D., 2011. Occurrence, partition and removal of pharmaceuticals in sewage water and sludge during wastewater treatment. Water Res. 45, 1165-1176. https://doi.org/10.1016/j.watres.2010.11.010

Jones-Lepp, T.L., Stevens, R., 2007. Pharmaceuticals and personal care products in biosolids/sewage sludge: the interface between analytical chemistry and regulation. Anal. Bioanal. Chem. 387, 1173-1183. https://doi.org/10.1007/s00216-006-0942-z

Keller, V.D.J., Williams, R.J., Lofthouse, C., Johnson, A.C., 2014. Worldwide estimation of river concentrations of any chemical originating from sewage-treatment plants using dilution factors. Environ. Toxicol. Chem. 33, 447-452. https://doi.org/10.1002/etc.2441

Knight, G.M., Glover, R.E., McQuaid, C.F., Olaru, I.D., Gallandat, K., Leclerc, Q.J., Fuller, N.M., Willcocks, S.J., Hasan, R., van Kleef, E., Chandler, C.I., 2021. Antimicrobial resistance and COVID-19: Intersections and implications. eLife 10, e64139. https://doi.org/10.7554/eLife.64139

Köck-Schulmeyer, M., Villagrasa, M., López de Alda, M., Céspedes-Sánchez, R., Ventura, F., Barceló, D., 2013. Occurrence and behavior of pesticides in wastewater treatment plants and their environmental impact. Sci. Total Environ. 458-460, 466-476. https://doi.org/10.1016/j.scitotenv.2013.04.010

Kovalova, L., Siegrist, H., Singer, H., Wittmer, A., McArdell, C.S., 2012. Hospital Wastewater Treatment by Membrane Bioreactor: Performance and Efficiency for Organic Micropollutant Elimination. Environ. Sci. Technol. 46, 1536-1545. https://doi.org/10.1021/es203495d

Krentz, A.J., Bailey, C.J., 2005. Oral Antidiabetic Agents. Drugs 65, 385-411. https://doi.org/10.2165/00003495-200565030-00005

Kuroda, K., Li, C., Dhangar, K., Kumar, M., 2021. Predicted occurrence, ecotoxicological risk and environmentally acquired resistance of antiviral drugs associated with COVID-19 in environmental waters. Sci. Total Environ. 776, 145740. https://doi.org/10.1016/j.scitotenv.2021.145740

Le Corre, K.S., Ort, C., Kateley, D., Allen, B., Escher, B.I., Keller, J., 2012. Consumption-based approach for assessing the contribution of hospitals towards the load of pharmaceutical residues in municipal wastewater. Environ. Int. 45, 99-111. https://doi.org/10.1016/j.envint.2012.03.008

Li, Z., Undeman, E., Papa, E., S. McLachlan, M., 2018. High-throughput evaluation of organic contaminant removal efficiency in a wastewater treatment plant using direct injection UHPLC-Orbitrap-MS/MS. Environ. Sci. Process. Impacts 20, 561-571. https://doi.org/10.1039/C7EM00552K

López Rocha, C.J., Álvarez-Castillo, E., Estrada Yáñez, M.R., Bengoechea, C., Guerrero, A., Orta Ledesma, M.T., 2020. Development of bioplastics from a microalgae consortium from wastewater. J. Environ. Manage. 263, 110353. https://doi.org/10.1016/j.jenvman.2020.110353

Lopez-Herguedas, N., González-Gaya, B., Castelblanco-Boyacá, N., Rico, A., Etxebarria, N., Olivares, M., Prieto, A., Zuloaga, O., 2022. Characterization of the contamination fingerprint of wastewater treatment plant effluents in the Henares River Basin (central Spain) based on target and suspect screening analysis. Sci. Total Environ. 806, 151262. https://doi.org/10.1016/j.scitotenv.2021.151262

Mansouri, K., Grulke, C.M., Judson, R.S., Williams, A.J., 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J. Cheminformatics 10, 10. https://doi.org/10.1186/s13321-018-0263-1

Markert, N., Rhiem, S., Trimborn, M., Guhl, B., 2020. Mixture toxicity in the Erft River: assessment of ecological risks and toxicity drivers. Environ. Sci. Eur. 32, 51. https://doi.org/10.1186/s12302-020-00326-5

Martín, J., Camacho-Muñoz, D., Santos, J.L., Aparicio, I., Alonso, E., 2012. Occurrence of pharmaceutical compounds in wastewater and sludge from wastewater treatment plants: Removal and ecotoxicological impact of wastewater discharges and sludge disposal. J. Hazard. Mater., Occurrence and fate of emerging contaminants in municipal wastewater treatment systems 239-240, 40-47. https://doi.org/10.1016/j.jhazmat.2012.04.068

Montes, R., Rodil, R., Rico, A., Cela, R., González-Mariño, I., Hernández, F., Bijlsma, L., Celma, A., Picó, Y., Andreu, V., de Alda, M.L., López-García, E., Postigo, C., Pocurull, E., Marcé,

R.M., Rosende, M., Olivares, M., Valcárcel, Y., Quintana, J.B., 2020. First nation-wide estimation of tobacco consumption in Spain using wastewater-based epidemiology. Sci. Total Environ. 741, 140384. https://doi.org/10.1016/j.scitotenv.2020.140384

Nannou, C., Ofrydopoulou, A., Evgenidou, E., Heath, D., Heath, E., Lambropoulou, D., 2020. Antiviral drugs in aquatic environment and wastewater treatment plants: A review on occurrence, fate, removal and ecotoxicity. Sci. Total Environ. 699, 134322. https://doi.org/10.1016/j.scitotenv.2019.134322

Nason, S.L., Lin, E., Eitzer, B., Koelmel, J., Peccia, J., 2022. Changes in Sewage Sludge Chemical Signatures During a COVID-19 Community Lockdown, Part 1: Traffic, Drugs, Mental Health, and Disinfectants. Environ. Toxicol. Chem. 41, 1179-1192. https://doi.org/10.1002/etc.5217

Nilsen, E., Smalling, K.L., Ahrens, L., Gros, M., Miglioranza, K.S.B., Picó, Y., Schoenfuss, H.L., 2019. Critical review: Grand challenges in assessing the adverse effects of contaminants of emerging concern on aquatic food webs. Environ. Toxicol. Chem. 38, 46-60. https://doi.org/10.1002/etc.4290

Osborne, V., Davies, M., Lane, S., Evans, A., Denyer, J., Dhanda, S., Roy, D., Shakir, S., 2020. Lopinavir-Ritonavir in the Treatment of COVID-19: A Dynamic Systematic Benefit-Risk Assessment. Drug Saf. 43, 809-821. https://doi.org/10.1007/s40264-020-00966-9

Perkons, I., Tomsone, L.E., Sukajeva, V., Neilands, R., Kokina, K., Pugajeva, I., 2022. Qualitative fingerprinting of psychoactive pharmaceuticals, illicit drugs, and related human metabolites in wastewater: A year-long study from Riga, Latvia. J. Environ. Chem. Eng. 10, 108110. https://doi.org/10.1016/j.jece.2022.108110

Qi, W., Singer, H., Berg, M., Müller, B., Pernet-Coudrier, B., Liu, H., Qu, J., 2015. Chemosphere 119, 1054-1061.

Reinstadler, V., Ausweger, V., Grabher, A.-L., Kreidl, M., Huber, S., Grander, J., Haslacher, S., Singer, K., Schlapp-Hackl, M., Sorg, M., Erber, H., Oberacher, H., 2021. Monitoring drug consumption in Innsbruck during coronavirus disease 2019 (COVID-19) lockdown by wastewater analysis. Sci. Total Environ. 757, 144006. https://doi.org/10.1016/j.scitotenv.2020.144006

Roveri, V., Lopes Guimarães, L., 2023. In silico prediction of persistent, mobile, and toxic pharmaceuticals (PMT): A case study in São Paulo Metropolitan Region, Brazil. Comput. Toxicol. 25, 100254. https://doi.org/10.1016/j.comtox.2022.100254

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ. Sci. Technol. 48, 2097-2098. https://doi.org/10.1021/es5002105

Singh, S., Roy, D., Sinha, K., Parveen, S., Sharma, G., Joshi, G., 2020. Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. Psychiatry Res. 293, 113429. https://doi.org/10.1016/j.psychres.2020.113429

Solaun, O., Rodríguez, J.G., Menchaca, I., López-García, E., Martínez, E., Zonja, B., Postigo, C., López de Alda, M., Barceló, D., Borja, Á., Manzanos, A., Larreta, J., 2021. Contaminants of emerging concern in the Basque coast (N Spain): Occurrence and risk assessment for a better monitoring and management decisions. Sci. Total Environ. 765, 142765. https://doi.org/10.1016/j.scitotenv.2020.142765

Usman, M., Farooq, M., Hanna, K., 2020. Environmental side effects of the injudicious use of antimicrobials in the era of COVID-19. Sci. Total Environ. 745, 141053. https://doi.org/10.1016/j.scitotenv.2020.141053

Wang, S., C. Green, H., L. Wilder, M., Du, Q., L. Kmush, B., B. Collins, M., A. Larsen, D., Zeng, T., 2020. High-throughput wastewater analysis for substance use assessment in central New York during the COVID-19 pandemic. Environ. Sci. Process. Impacts 22, 2147-2161. https://doi.org/10.1039/D0EM00377H

WHO, 2021. COVID-19 weekly epidemiological update. World Health Organization.

Bilde: Marco Verch

Gazteak eta COVID-19aren txertoarekiko jarrera: adinak eta gurasoen ikasketa-mailak baldintzatzen dute haien iritzia

Birus baten pandemiaren erdian, populazioaren txertaketa orokorra bilakatu da talde-immunitatea lortu eta pandemia amaitzeko estrategiarik eraginkorrena. Horretarako, beharrezkoa da erresistentzia erakusten duten giza taldeak identifikatzea eta haien kezkei erantzun bat ematea. Gazteen artean, gazteenek eta ikasketa-maila baxueneko gurasoak dituztenek erakutsi dute txertoekiko erresistentzia handiena. Ezinbestekoa da gazteen erresistentziak kontuan hartzea kanpaina publikoak diseinatzeko garaian.

Pandemian zehar, behin baino gehiagotan ikertu da helduek txertoekiko zer-nolako jarrera duten, baina apenas egon da daturik gazte eta haurren iritziaz. Helduei galdetu zaie seme-alabak txertatzeko asmoaren inguruan, baina ez da aztertu nerabeen beren iritzia. Eta bada adin-tarte zabal bat non, gurasoen zaintzapean izanik ere, iritzi propioa erakuts baitezakete eta erabakiaren parte izan. Beraz, garrantzitsua da haien iritzia aztertzea. Batez ere, haur eta nerabeen eskolaratzea eta txertaketa oso eztabaidatuak izan direlako pandemian zehar. Eta txertoak 16 urtetik beherako gazteentzat beranduago baimendu izanak hutsune bat sortu duelako immunizatutako populazioan, eta talde-immunitatea lortzea eragotzi.

Gazteen iritzia berariaz aztertu denean, ikusi da % 68,3k txertatzeko asmoa zutela, % 15 zalantzan zeudela eta % 15ek ez zutela txertoa hartu nahi. Datuek txertoekiko jarrera orokor ona uzten dute agerian, helduen txertatzearen aldeko tasaren (% 70) oso antzekoa.

Adina eta gurasoen hezkuntza-maila

Gazteen iritzian ez dute generoaren araberako desberdintasunik ikusi, baina bai adinaren araberakoa: gazteenek dute txertoarekiko errezelorik handiena. Eta bigarren aldagai garrantzitsu bat nabarmendu da: gurasoen hezkuntza-mailak zuzenean eragiten du gazteen txertoarekiko iritzian. Guraso bietatik bakarrak ere unibertsitate-mailako ikasketarik ez duten gazteek erakutsi dute errezelorik handiena. Ikasleek emandako datuen arabera, % 26k ez dute unibertsitate-mailako hezkuntza duen gurasorik, % 29,8k bakarra dute eta % 28,8k bi gurasoak.

Bestetik, gazteen % 15k aitortu dute gaitz kroniko bat duela —asma, alergia, arreta-defizitak eragindako nahasmendua eta hiperaktibitatea, besteak beste—, eta horrek txertoekiko iritzia baldintzatzen ote duen jakin nahi izan da, talde zaurgarri moduan identifikatu izan baitira. Baina ez da desberdintasun esanguratsurik ikusi batzuen eta besteen artean.

Beste ikerketa batzuen arabera, txertoen segurtasuna, balizko albo-ondorioak eta beste pertsona batzuek txerto hori gehiago behar dutela pentsatzea dira zalantza-eragilerik handienak txertoa hartzeari dagokionez. Txertoa jasotzearen aurrean zalantza erakutsi duten % 15 horiek dira informazio-kanpaina publikoen itu nagusia, gerta bailiteke txertoa hartzearen alde egitea informazio egokia emanez gero eta lasaitasuna transmitituz gero.

Garrantzitsua da txertoa hartzeko errezelo handiena erakusten duten taldeak identifikatzea eta haien kezkei erantzun bat ematera bideratzea kontzientziazio-kanpaina publikoak.

InpAkta 00 (2023) 150-157

https://doi.org/10.1007/s00431-021-04343-1

Students' age and parental level of education influence COVID-19 vaccination hesitancy

Anna Zychlinsky Scharff[1*], Mira Paulsen[2], Paula Schaefer[3], Fatma Tanisik[3], Rizky Indrameikha Sugianto[2], Nils Stanislawski[4], Holger Blume[4], Bernhard M. W. Schmidt[5], Stefanie Heiden[6], Meike Stiesch[3], Anette Melk[2]

[1] Dept. of Pediatric Hematology and Oncology, Hannover Medical School; [2] Dept. of Pediatric Kidney, Liver, and Metabolic Diseases, Hannover Medical School; [3] Dept. of Prosthetic Dentistry and Biomedical Material Research, Hannover Medical School; [4] Institute of Microelectronic Systems, Leibniz University Hanover; [5] Dept. of Nephrology and Hypertension, Hannover Medical School; [6] Institute of Innovation Research, Technology Management & Entrepreneurship, Leibniz University Hanover

ABSTRACT: Widespread vaccination in pursuit of herd immunity has been recognized as the most promising approach to ending the global pandemic of coronavirus disease 19 (COVID-19). The vaccination of children and adolescents has been extensively debated and the first COVID-19 vaccine is now approved in European countries for children aged > 12 years of age. Our study investigates vaccination hesitancy in a cohort of German secondary school students. We assessed 903 students between age 9 and 20 in the period between 17 May 2021 and 30 June 2021. 68.3% (n = 617) reported intention to undergo COVID-19 vaccination, while 7% (n = 62) did not want to receive the vaccine and 15% (n = 135) were not yet certain. Age and parental level of education influenced COVID-19 vaccine hesitancy. Children under the age of 16 as well as students whose parents had lower education levels showed significantly higher vaccine hesitancy.

Conclusion: Identifying subsets with higher vaccination hesitancy is important for targeting public information campaigns in support of immunization.

What is Known:

- The willingness to receive COVID-19 vaccination among adults in Europe is about 70%, but data for children and adolescents is lacking.
- The lack of immunization in younger cohorts represents a significant barrier to achieving herd immunity, and also leaves children and adolescents vulnerable to acute and long-term morbidity from natural COVID-19 infections.

What is New:

- Intention-to-vaccinate among children and adolescents is high (~ 70%); conversely, vaccination hesitancy is low.
- Age and parental level of education influenced COVID-19 vaccine hesitancy among children and adolescents.

* **Corresponding author / Harremanetan jartzeko:** Anna Zychlinsky Scharff, Department of Pediatric Hematology and Oncology, Hannover Medical School, Hannover, Germany, Spain. e-mail: zychlinskyscharff.anna@mh-hannover.de. **Contributed by** Maria Napal Fraile **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Zychlinsky Scharff, Anna *et al.* «Students' age and parental level of education influence COVID-19 vaccination hesitancy», *European Journal of Pediatrics* (2022) 181:1757–1762 (https://doi.org/10.1007/s00431-021-04343-1)

1. Introduction

Since the development of novel vaccines against coronavirus disease 19 (COVID-19), widespread vaccination in pursuit of herd immunity has become the most promising path to end the pandemic [1]. Most COVID-19 vaccines are approved exclusively for adults. The BNT162b2 COVID-19 vaccine, produced by BioNTech/Pfizer, was initially approved for individuals over the age of 16 and this was expanded to those over 12 years old in June 2021, though the recommendations by the German vaccine regulating agency were more restrictive [2]. The lack of immunization in younger cohorts represents a significant barrier to achieving herd immunity, and leaves children and adolescents vulnerable to acute and long-term morbidity from natural COVID-19 infections [3]. While children represent a minority of severe disease courses and deaths in comparison to adults, they carry a significant burden of disease [4]. Children with underlying conditions and infants under 1 year of age are at risk for severe initial infection [4], and children of all ages, can develop multi-system inflammatory disease in children (MIS-C, also known as PIMS-TS), even after asymptomatic or pauci-symptomatic COVID-19 [5].

As trials demonstrating the safety and efficacy of COVID-19 vaccines in pediatric populations emerge, policy makers worldwide are embarking on campaigns to vaccinate their youngest citizens, making the question of vaccine hesitancy in this group more relevant than ever. Germany, like other European governments, announced on 11 May 2021 that it intends to offer vaccinations to 12- to 17-year olds by the end of the summer school break. Studies examining COVID-19 vaccination intention and hesitancy in this context have focused on parental views, as parents and guardians generally make medical decisions for minors under their care [6, 7]. However, the independent views of adolescents themselves have not been assessed, and this represents a significant informational void in the quest to vaccinate and protect pediatric populations.

In this study, we show that older children, those for whom the vaccine was approved from the initial launch, have higher rates of intention-to-vaccinate than younger ones. In addition, we observed that a lower parental educational level correlated to higher vaccine hesitancy. Conversely, children and adolescents whose parents attained higher education levels were more likely to report that they intended to get the COVID-19 vaccine.

2. Methods

2.1. Study design

Data was collected at the final time point of the longitudinal Transmission Analytic COVID-19 (TRAC-19) study, which explored SARS-CoV-2 infections and behavioral patterns in two secondary schools in Hannover, Germany [8]. Nine hundred thirteen students participated in the TRAC-19 study between May 17 and June 30, 2021, and 903 returned the questionnaires about vaccination hesitancy (Suppl. Table 1). In the context of the larger study, students provided nasal swap and blood sample, which were tested for SARS-CoV-2-specific antibodies by Elecsys® Anti-SARS-CoV-2 (Roche) assay according to the manufacturer's instruction. Venipuncture was completely optional; students could opt-out from blood sampling and still participate in the study.

Table 1
Logistic regression model for intention-to-vaccinate

Covariate	OR	95% CI		*p*-value
Sex				
Male vs. female	1.07	0.74	1.55	0.72
Age				
9-12 vs. 13-15 years	0.46	0.28	0.75	0.0005
9-12 vs. 16-19 years	0.29	0.15	0.54	< 0.0001
13-15 vs. 16-19 years	0.63	0.33	1.19	0.22
Chronic disease				
No vs. yes	0.51	0.23	1.14	0.10
College-educated adults				
No vs. one	0.42	0.25	0.72	0.0003
No vs. two	0.26	0.15	0.46	< 0.0001
One vs. two	0.61	0.33	1.11	0.13

OR, odds ratio; 95% CI, confidence interval.

The TRAC-19 study was approved by the institutional review board (No. 9085_BO_S_2020) and complies with the Declaration of Helsinki. Study participation was voluntary and informed consent was obtained from participants and, in case of minors, their legal guardian.

3. Statistics

Data are given as mean and standard deviation (SD) or numbers (*n*) and percentages (%). A multivariable logistic regression model was employed for vaccination hesitancy and was adjusted for multiple comparison with Šidák. For the analysis, two categories were built (yes vs. no/unsure). Covariates included general demographics (sex, age), chronic diseases, and parent's educational level (one, two, or no adults with college education). Age groups were built according to the categories defined by federal vaccine recommendations: 9-12 years, 13-15 years, and 16-19 years. *P*-values < 0.05 were considered significant. Statistical analyses were performed using SAS 9.4 (SAS Institute Inc., USA).

4. Results

A total of 903 students from grades 5 to 13 participated in this study. The mean age was 14.6 years (SD 2.3), and 52.4% were female (Suppl. Table 2). Fifteen percent (*n* = 134) reported having a chronic disease, most commonly asthma, allergies, or attention deficit hyperactivity disorder. Twenty-seven percent (*n* = 246) had no college-educated adult living in their household, 29.8% (*n* = 269) had one, and 28.8% (*n* = 260) had two college-educated adults. Fourteen percent (*n* = 130) of students did not provide sufficient information to determine parental education level and were therefore not categorized in any of the three groups.

Three percent of students (*n* = 28) reported having had COVID-19 (confirmed by polymerase chain reaction or antigen test). We detected SARS-CoV-2 antibodies in 17 of these students and in another 11 students that had not reported a previous positive test. Ten percent (*n* = 89) of students had already received at least one vaccine dose with the majority (*n* = 80) being older than 16 years and thus eligible for the vaccine from the initial release (Fig. 1a). Nine students had received the vaccine despite being below the age of eligibility at the time. Twenty-six of these early vaccinees reported having a chronic disease, accounting for 19.1% (26/136) of all students with chronic diseases.

A total of 903 students (99%) answered the questionnaire about vaccination hesitancy. In addition to those already vaccinated, 68.3% (*n* = 617) reported their intent to undergo COVID-19 vaccination (Fig. 1a). Seven percent (*n* = 62) did not want to receive the vaccine and 15% (*n* = 135) were still uncertain.

We performed a mixed model analysis to identify factors that influence students' intention to receive the vaccine (Table 1). We observed no sex difference but found differences between the age groups. Older students (age 13-19) showed significantly higher intention-to-vaccinate compared to the younger age group (Fig. 1b, Table 1). Notably, most of the students already vaccinated were adolescents (age 16-19). Students who declined vaccination or were uncertain were likely to belong to the younger age groups (Fig. 1b). To explore the influence of parental education on students' vaccine preferences, we included the educational status of adults living in the same household, presumably parents or guardians, in the model. Vaccination hesitancy was higher in households with no college- educated adults than in those with one or two college- educated adults (Table 1, Suppl. Table 3). We did not find a significant difference in students' intention-to-vaccinate between healthy individuals and those reporting chronic condition.

5. Discussion

Widespread vaccination is indispensable to ending the pandemic spread of COVID-19 [1]. As data on children and adolescents is lacking, we assessed vaccination willingness or hesitancy among a co-

hort of 903 students in May and June 2021, shor tly after vaccine access was expanded to include adolescents. Overall, the intention- to-vaccinate was high in our cohort of students, mirroring data from adults in Europe showing similar vaccine acceptance [9].

The vaccination of children and adolescents has been hotly debated, especially in the context of in-person learning. While the European Medicines Agency has already approved the BNT162b2 vaccine for children over 12 years of age in May 28, the Standing Committee on Vaccination (STIKO) in Germany has offered more restrictive only for those at high risk for severe COVID-19 [2], and finally extending the recommendation for all people aged 12 years or above in August 19. This discrepancy likely contributed to the age difference in vaccination hesitancy we observed, as access to vaccination may have affected willingness. Those students whose age placed them squarely in the eligible range according to the initial approval for BNT162b2 were more likely to report an intention to receive the vaccine. Indeed, nearly a third of students in this category had already received the first dose. Students in the 12-16 age group, for whom a vaccination was recommended by the STIKO only in the case of chronic illness during the study period, had higher rates of vaccine hesitancy, suggesting that timing of recommendations and the discrepancy between European and national guidelines may have influenced this cohort. Even though the STIKO vaccination guidelines changed only for children aged 12-16 years with a chronic condition during the observation period, this remains an important limitation to this study. Vaccination hesitancy is often highest in the initial phase of vaccination campaigns and may diminish as time goes on, and widespread vaccination contributes to an aura of normalcy around the newly introduced substance [10].

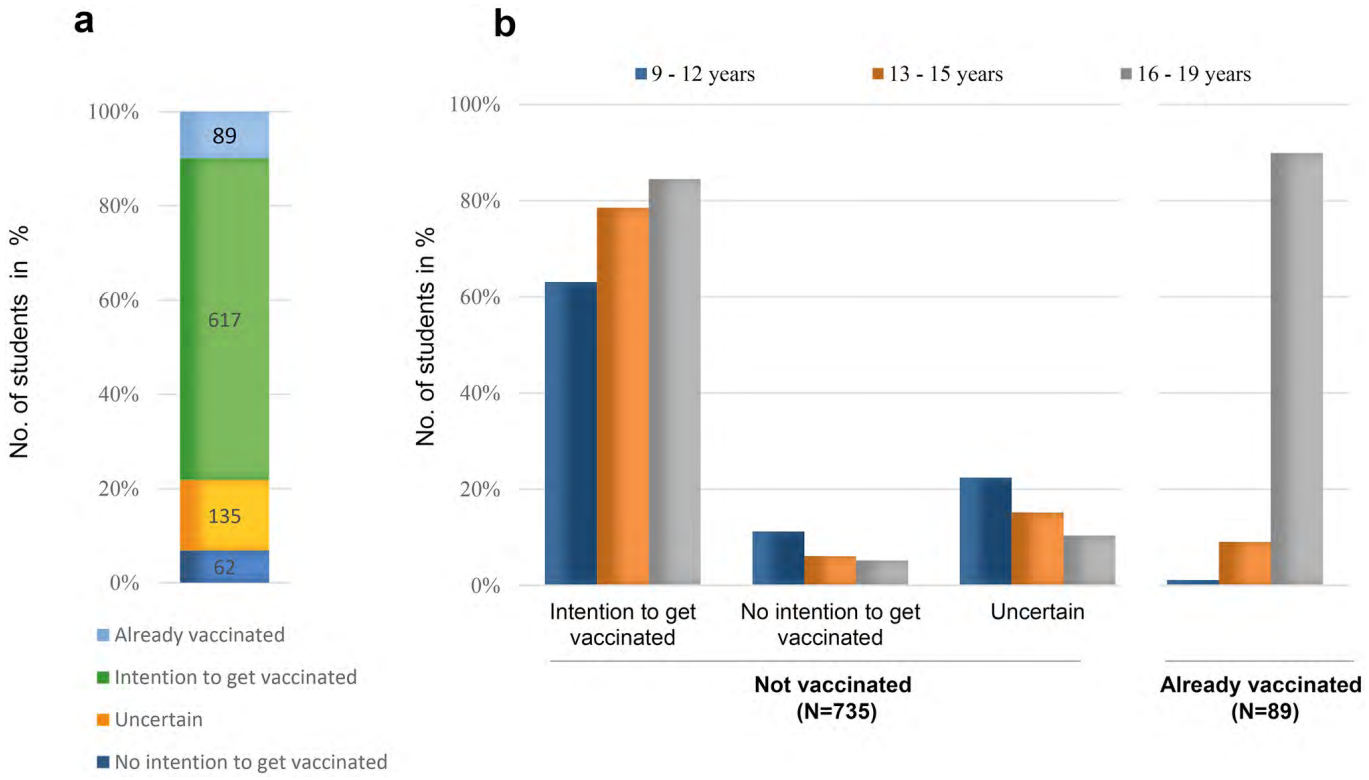


Figure 1. COVID-19 vaccination hesitancy. The intention to undergo COVID-19 vaccination for all students a and by age groups for those not vaccinated yet and students already vaccinated (b)

In our cohort, intention-to-vaccinate was higher in children and adolescents from households with college-educated caregivers. This is consistent with data showing that an educational level below a bachelor's degree predicted hesitancy towards routine childhood vaccinations and annual influenza vaccines [11]. Brandstetter *et al.* suggest that parental educational level influences parents' intention-to-vaccinate their children against COVID-19. However, this study only surveyed parents of younger children (1 to 5 years of age) examining a context in which health care choices are presumably predicated on parental preferences alone, and thus did not address the views of minor patients themselves [12]. We sought to examine an age group that, while still under the legal custody of parents or guardians, is likely to participate in decision-making around their own medical care and whose viewpoints and intentions are thus highly relevant. It is important to note, however, that German law, like in most European countries, requires parental consent for vaccination in individuals under the age of 18 years. While this limits the agency of minors to some extent, a German court ruled that minors may receive a COVID-19 vaccination even against the wishes of one parent, provided the vaccination is in accordance with STIKO guidelines and the minor and the other parent agrees to vaccination [13]. It is interesting to note in the context of Brandstetter *et al.* [12] that parental education level seems to affect not only parents' attitude towards COVID-19 vaccines, but also that of their offspring.

Importantly, 15% of our cohort was still uncertain about receiving a COVID-19 vaccine. This hesitant but likely convincible subset is the most relevant target group of public information campaigns in support of immunization. While our study did not explore the reasons underlying the reported uncertainty, previous studies suggest that vaccination safety, concerns over side effects, and believing others in greater need of the vaccine may play a role [7, 14].

6. Conclusion

Overall, we report high levels of vaccine willingness, although younger age and lower parental education levels correlated with higher vaccine hesitancy. Identifying subsets with a higher vaccination hesitancy is important for the targeting of public information campaigns in support of immunization and achieving herd immunity.

6.1. *Supplementary information*

The online version contains supplementary material available at https://doi.org/10.1007/s00431-021-04343-1.

6.2. *Acknowledgements*

We gratefully acknowledge the contributions of Beate Günther, Philipp Tups, Katharina Kalinowski, Rita Wonik-Schmidt, Brigitte Naber, and Frank Weinberg for their organizational support at the schools. We would like to thank all study participants, including the students, teachers, and staff at the Schillerschule and IGS Roderbruch for their support of the TRAC-19 study.

6.3. *Authors' contributions*

AZS and MP conceptualized and designed the study, drafted the initial manuscript, carried out initial analyses, and contributed to the interpretation of data. PS and FT collected data and coordinated and supervised data collection. RS performed the statistical analyses and contributed to the interpretation of data. NS and HB designed the data collection instruments and software, coordinated and supervised data collection, and carried out initial analyses. AM, BS, MS, and SH conceptualized and designed the study, and critically reviewed and revised the manuscript. All authors provided input critically reviewed and approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

6.4. *Funding*

Open Access funding enabled and organized by Projekt DEAL. The study, as part of the larger TRAC-19 study, is funded by the Ministry of Science and Culture of Lower Saxony, Germany, reference number 14-76103-184. The Ministry of Science and Culture

had no role in the design and conduct of the study; data collection, management, analysis or interpretation; or writing of the report.

6.5. Availability of data and material

De-identified patient datasets will be available upon written request to the corresponding author following publication.

7. Declarations

Ethics approval

The study was approved by the institutional ethical committee.

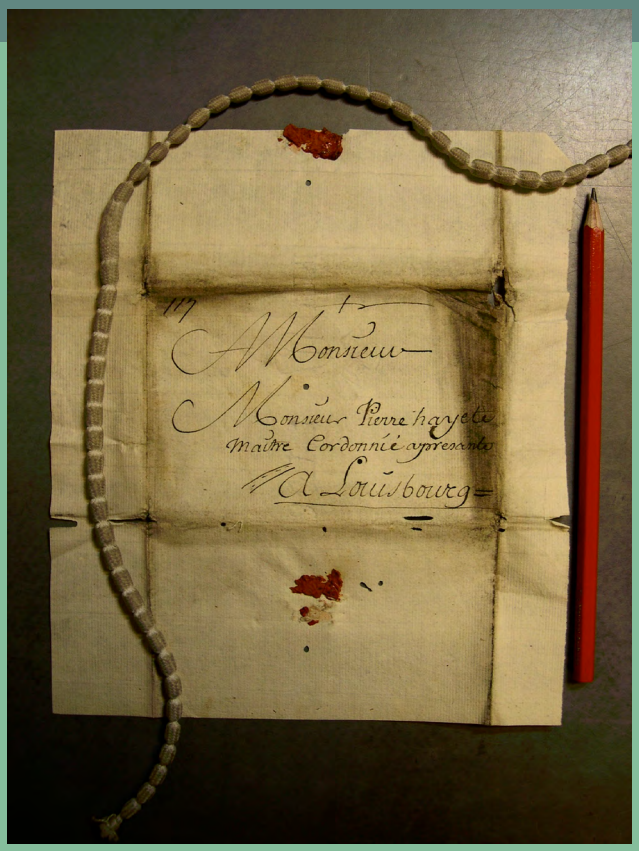
7.1. Consent to participate

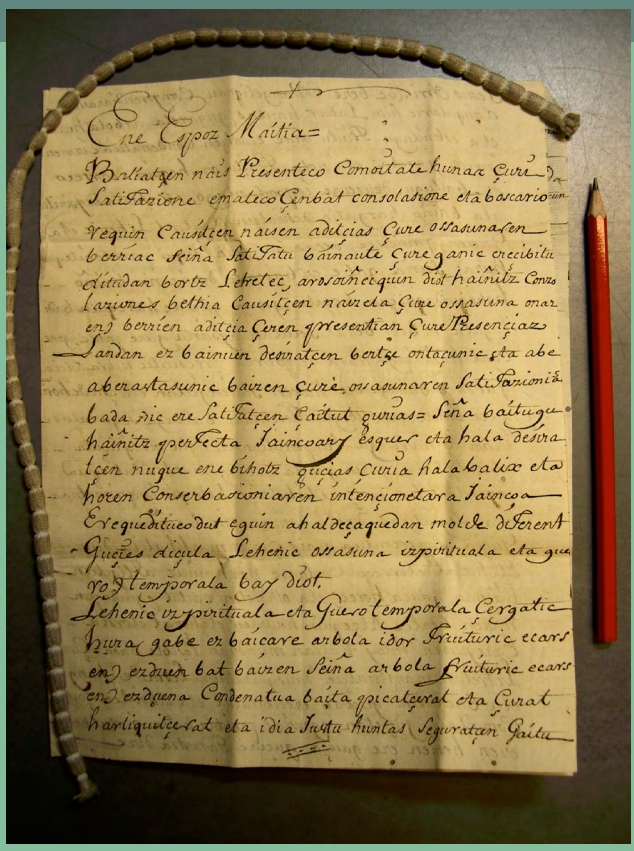
All participants gave their written informed consent prior to their enrolment in the study.

8. References

1. World Health Organization (2021) COVID-19 Strategic preparedness and response plan. Geneva. Licence: CC BY-NC-SA 3.0 IGO
2. Federal ministry of health (2021) Vaccines against COVID-19: Comirnaty® from BioNTech/Pfizer. Germany. https://www.zusammengegencorona.de/impfen/impfstoffe/impfstoffe-gegen-covid-19-comirnaty-r-von-biontech-pfizer/. Accessed 31 Oct 2021
3. Velavan TP, Pollard AJ, Kremsner PG (2020) Herd immunity and vaccination of children for COVID-19. Int J Infect Dis 98:14-15. https://doi.org/10.1016/j.ijid.2020.06.065
4. Bellino S, Punzo O, Rota MC, Del Manso M, Urdiales AM, Andrianou X, Fabiani M, Boros S, Vescio F, Riccardo F, Bella A (2020) COVID-19 disease severity risk factors for pediatric patients in Italy. Pediatrics 146(4). https://doi.org/10.1542/peds.2020-009399
5. Verdoni L, Mazza A, Gervasoni A, Martelli L, Ruggeri M, Ciuffreda M, Bonanomi E, D'Antiga L (2020) An outbreak of severe Kawasaki- like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. The Lancet 395:1771-1778. https://doi. org/10.1016/S0140-6736(20)31103-X
6. Montalti M, Rallo F, Guaraldi F, Bartoli L, Po G, Stillo M, Perrone P, Squillace L, Dallolio L, Pandolfi P, Resi D, Fantini MP, Reno C, Gori D (2021) Would parents get their children vaccinated against SARS-CoV-2? Rate and predictors of vaccine hesitancy according to a survey over 5000 families from Bologna, Italy. Vaccines 9(4):366. https://doi.org/10.3390/vaccines9040366
7. Ruggiero KM, Wong J, Sweeney CF, Avola A, Auger A, Macaluso M, Reidy P (2021) Parents' intentions to vaccinate their children against COVID-19. J Pediatr Health Care. https://doi.org/10.1016/j.pedhc.2021.04.005
8. Paulsen M, Scharff AZ, de Cassan K, Sugianto RI, Blume C, Blume H, … Melk A (2021) Children and adolescents' behavioral patterns in response to escalating COVID-19 restriction reveal sex and age differences. J Adolesc Health. https://doi.org/10.1016/j. jadohealth.2021.11.021
9. Neumann-Bohme S, Varghese NE, Sabat I, Barros PP, Brouwer W, van Exel J, Schreyogg J, Stargardt T (2020) Once we have it, will we use it? A European survey on willingness to be vaccinated against COVID-19. Eur J Health Econ 21:977-982. https://doi.org/10.1007/s10198-020-01208-6
10. Hanson KE, Koch B, Bonner K, McRee AL, Basta NE (2018) National trends in parental human papillomavirus vaccination intentions and reasons for hesitancy, 2010-2015. Clin Infect Dis 67:1018-1026. https://doi.org/10.1093/cid/ciy232
11. Kempe A, Saville AW, Albertin C, Zimet G, Breck A, Helmkamp L, Vangala S, Dickinson LM, Rand C, Humiston S, Szilagyi PG (2020) Parental hesitancy about routine childhood and influenza vaccinations: a national survey. Pediatrics 146(1). https://doi.org/10.1542/peds.2019-3852
12. Brandstetter S, Bohmer MM, Pawellek M, Seelbach-Gobel B, Melter M, Kabesch M, Apfelbacher C, group KU-Ks (2021) Parents' intention to get vaccinated and to have their child vaccinated against COVID-19: cross-sectional analyses using data from the KUNO-Kids health study. Eur J Pediatr. https://doi.org/10.1007/ s00431-021-04094-z
13. Legal Tribune Online (2021) Higher Regional Court of Frankfurt a.M. on parental consent to vaccination: the one decides, who follows the STIKO. https://www.lto.de/persistent/a_id/45810/. Accessed 16 Oct 2021
14. Adams SH, Schaub JP, Nagata JM, Park MJ, Brindis CD, Irwin Jr CE (2021) Young adult perspectives on COVID-19 vaccinations. J Adolesc Health. https://doi.org/10.1016/j.jadohealth.2021.06.003

8.1. Publisher's Note





«Lettre de Francha Barrere à Pierre Hayete», Entrepôt de données ANPERSANA, https://anpersana.iker.univ-pau.fr/items/show/131

Le Dauphin ontziko gutunak ikertzea, ezinbestekoa XVIII. mendeko euskara eta euskal gizartea ezagutzeko

Xabier Lamikiz ikertzaileak euskaraz idatzitako 47 gutun aurkitu zituen Londresko The National Archives-en, 2003an. 1757an Lapurdiko herritarrek Kanadako Louisbourge herrian bizi ziren senideei idatzitako gutunak dira, *Le Dauphin* ontziak zeramatzanak. Dokumentu historiko baliotsuak dira garai hartako Lapurdiko euskara eta gizartea ezagutzeko.

1757ko apirilaren 9an, Britainia Handiko hiru gerra-ontzik *Le Dauphin* itsasontzi frantsesa harrapatu zuten. Gerra betean, Baionatik atera eta Atlantikoa zeharkatzen ari zen, Kanadarako bidean. Itsasontzia ez ezik, barnean zeraman guztia bahitu zuten britainiarrek, dokumentazioa barne. Dokumentuon artean euskaraz idatzitako 50 eskutitz zeuden, Lapurdiko 48 bizilagunek Kanadako Louisbourg herrian zeuden 45 pertsonari idatziak.

«1757an euskaraz idatzitako gutunak dokumentu historiko baliotsuak dira, garai hartako euskara jasoa zein soilaren ezaugarriak ezagutzeko»

Kortsario-ontzia izanik ere, merkataritza-lanak egiten ari zen bidaia hartan *Le Dauphin*. Donibane Lohizune, Azkaine, Urruña, Ziburu, Senpere eta inguruko beste herri batzuetako marinelak ziren gehienak. Gutunak Lapurdiko 35 emakumek eta 12 gizonek idatzi zituzten, eta hartzaileak Euskal Herritik Louisbourgera joandako gizonezkoak ziren, batez ere. Semeari, senarrari, anaiari zein adiskideei idatzitako gutunak dira gehienak. Besteak beste, herriko albisteak emateko, heriotzak jakinarazteko, gerraren berri emateko, zerbait debekatzeko edo baimena emateko idazten zieten. Baina haserrea ere adierazten zuten askok, eta kexu ziren laguntza ekonomikorik ez bidaltzeaz, «beste aldean» amodioak izateaz eta gutunek erantzunik ez jasotzeaz.

Informazio baliotsua eskura

Ikertzaileen ustez, garai hartako euskara aztertzeko baliabide baliotsua da Londresko *The National Archives-en* aurkitutako dokumentazioa: euskararen ortografia, lexikoa, morfosintaxia, estiloa eta bestelako ezaugarriak aztertzeko aukera ematen du. Bestetik, garai hartako literaturan erabiltzen zen euskara jasoaren eta herritar soilek idazten zutenaren artean zenbateko aldea zegoen ikus daiteke. Nolabait, garai hartan zer nolako alfabetizazio-maila eta eskolaratzea zegoen ondorioztatu daiteke. Beraz, XVIII. mendeko Lapurdiko gizartea nolakoa zen jakiteko ere baliagarria da.

Londresko artxiboan oraindik aztertu gabe dagoen dokumentazioa begiratzea interesgarri jo dute ikertzaileek, informazio osagarria lor daitekeelakoan. Esaterako, Lapurdiko, Nafarroa Behereko eta Zuberoako herritarrek ez bezala, hegoaldeko probintzietako euskal herritarrek gaztelaniaz idazten zituzten gutunak, beren artean euskaraz mintzo ziren arren. Londresko dokumentazioa aztertzeak horri buruzko datu gehiago argitu ditzakeela uste dute ikertzaileek. Izan ere, 1652 eta 1815 bitartean konfiskatutako 160.000 gutun daude gordeta, eta, Europako hainbat hizkuntzatan daudenez idatziak dokumentuak, aberasgarria izan daiteke beste hizkuntzetako eta beste eremuetako idatziekin konparatzea.

InpAkta 00 (2023) 158-168

https://doi.org/10.1387/InpAkta.25019

Le Dauphin ontziaren euskarazko gutuneriaz (Baiona, 1757)

Xarles Videgain[1]

IKER UMR 5478 CNRS, UPPA

Londresen, TNA (*The National Archives*)-ko egoitzan, gaur arte atxikiak izan dira 1652tik 1815 arte eta 1854tik 1856 arte Angelesek presa gisa hartu zituzten ontzietako dokumentu aberatsak. Hain zuzen, arte horietan, Inglaterrako ontzidiak 36 000 barku harrapatu zituen. Bornuan zeuden 160 000 gutunak kontserbatuak izan dira, ikerlarien bazka. Europako hizkuntz hanitzetan datoz dokumentuak, eta parte bat euskaraz ere.

Xabier Lamikizi zor diogu altxor horren berri izatea. Serendipitateak deus guti izan zuen egitekorik haren lanean. Alabaina bere tesia prestatzen ari zelarik, dokumentazioari ximenki so eginez, iduriz besteak bezalako kaja bat ideki zuelarik, haren barneko gutunak eskuztatuz ohartu zen *Le Dauphin de Louisbourg* deitu ontzian zeuden gutunak Lapurdiko jendeek igorritakoak bazeudela Euskaraz idatziak ziren 49 gutun berezi zituen (gero beste bi agertuko ziren). Gutunak joan behar ziren ontsalaz L'Îsle Royaleko Louisbourg hirira, han bizi ziren Lapurtarrengana. Arraintzatik bizi ziren gehien bat, nahiz parte bat ofiziale, sehi ala mutil zen. X. Lamikiz berehala ohartu zen gutun horien balioaz eta Iker gaztigatu zuen. Ondotik ikerkideok, beste ikerle batzuek eta TNAko *Prize Letters* saileko arduradun den R. Cock jaunak euskarazko gutun eta dokumentu gehiago deskubritu dugu nahiz ez ditugun hemen aipatuko. X. Lamikizen tesiaren beraren ondotik ondoko lanak agertu ziren : Lamikiz X., Padilla-Moyano M. & Videgain Ch. 2015; Carette 2019; Lafuente A. 2021; Talec & Videgain (argitaratzekoa). CORBAN deitu proiektuan, Ph. Chareyre eta C. Mounole erakasleen zuzendaritzapean, tesia egiten ari da A. Lafuente (UPPA).

[1] *Le Dauphin* deitu proiektua CNRSko IKER-UMR-5478 laboratorioan eramana izan da « Bordeleko Maison des Sciences de l'Homme en Aquitaine » deitu egiturak sustengatu gaituelarik. Aldez edo moldez proiektua lagundu dutenei milesker : X. Lamikiz; Ur Apalategi; A. Arkotxa-Scarcia ; P. Baudry ; A. Bevan ; C. Carette ; M. C. Castérot ; Ph. Chareyre ; R. Cock ; G. Denis ; U. Etxebarria ; R. Etxepare ; A. Fuentes Zamalloa ; A. Lafuente ; M Lavergne ; C. Mounole ; I. Duguine ; M. Padilla ; J. Peytavi ; J. P. Talec ; A. Viaut.

* **Corresponding author / Harremanetan jartzeko:** Xarles Videgain. IKER UMR 5478 CNRS, UPPA. e-mail: charles.videgain@univ-pau.fr.
Contributed by Urtzi Etxeberria **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Videgain, Xarles. «*Le Dauphin* ontziaren euskarazko gutuneriaz (Baiona, 1757)», *InpAkta* 00 (2023) 160-168 (https://doi.org/10.1387/InpAkta.25019)

Historikoki, hastapeneko gertakaria hauxe izan zen: Baionatik 1757ko apirileko 4an atera zen *Le Dauphin de Louisbourg* deitu ontzia. 80 toneladako ontzia zen. Arrutagai zuen Canadako l'Île Royale, gaur *Cape Breton Island* izena daukan ugartea. Baionatik atera eta egun gutiren barne, apirileko 9an, ontzia harrapatu zuten *Somerset*, *Devonshire* eta *Rochester* deitu ontzi ingelesek (Cock, 2023). Porstmouthen ontziko arduradunen itaunketa apirileko 23an hasi zen : artxiboetan gelditu diren auziaren dokumentuak laster argitaratuko ditugu.

Bornuan 6 «*pieza*» (kanoiak) bazeuzkan ontziak zer gerta ere bere buruaren defenditzeko, edo parada izan balu bera baino ahulagoko ontzi baten presa egiteko. Beti ere, arma horiek ez ziren asko ezen bere kargamentuan merkantziak ere Louisbourgera eraman behar zituen *Le Dauphin*-ek: zapetak, aguardientea, naipe-jokoak, pakotillak eta beste. Horrengatik «Gerlako ta merkantzietako komisioa» firmatu zion 1757ko martxoko 29an Jean-Marie de Bourbon, Penthièvreko duke eta Frantziako almiranteak. Baimena hori ezinbestekoa zitzaion ontziari, Zazpi Urtetako Gerla hasiberria zelako.

Baiona 1757an zertan zen J. Pontetek argiki erakusten du. Portuan kortsuak eta itsasoko jardunaldiek zuten garrantzia azpimarratzen du. Ahatik, *Le Dauphin* ontzia ez zela kortsario bat zehazten du J. Pontetek. Bere 1690-ko hiztegian, Furetièrek «pirate» eta «corsaire» elkarretarik argiki berezten zituen. Pirata batek itsasoan gaindi jotzen zuen (du) Erregeren edo Estatuaren baimenik gabe; kortsarioa aldiz harrapakinen ondotik itsasoan bazebilen baina Erregek baimenduta eta Erregeri bere mozkinetik parte bat emanez. Egia da ez zuela besterik egiten, hots «full time» ari zen, ez zuen ez komertziorik ez garraiorik egiten. *Le Dauphin* aldiz ez zen egiazko kortsario bat bere karga baitzuen Canadan saltzekoa eta lehenik hara eramatekoa. Kanoiak balin bazituen ere, piezak ez ziren nahiko azkar kortsario huts gisa ibil zedin. J. Pontetek ohartarazi bezala, kortsua jardunaldi oso arautua zen eta Estaduak egin ez zezakeena kortsarioak hartzen zuen bere gain. Bestela erranez, kortsarioek osatzen zuten Estatuaren indarra, honek ordukotz ez baitzuen indarraren eta bortizkeriaren monopolioa segurtatzen ahal. Itsasoan presa egiteko eskubideaz luzaz mintza gintezke hemen, Hobbes (*Léviathan*), Grotius (*Mare Liberum, De Jure Belli ac Pacis*), edo goizago Francisco Vitoria bera (*De potestate civili, De Jure belli Hispanorum in barbaros*) aipatuz, herri katolikoek eta protestanteek izan zituzten gatazka guziekin.

Beti ere, Zazpi Urtetako Gerla denboran, Baiona zen Frantziako erresuman presa gehien egiten zuen portua: Dunkerquek eta Saint-Malok baino haboro. Jardunaldi eta ekonomikoa eta belikoa zen: zonbat diru ekartzen zuen zehatz kalkulatu du J. Pontetek: 15 milioi libera eman zuten presa guziek, horrek egin baitzuen 6 milioitako mozkina, Erregeren, burgesen (armadore), kapitainen eta mariñelen artean barreatu zena. Baina kortsuak bazuen itzalik: heriotzarena. Kortsuan eta gerlan hildakoen kopuru handia gertatu zen eta adibidez 485 lapurtar hil ziren Inglaterrako presondegi gogorretan. Poydenot, Baionako burgesak erraten zuen bezala, abantail deitoragarri eta zorigaiztokoa izaten zen kortsoa.

Le Dauphin ontziaz den bezenbatean, Jean Lanne zen armadore eta Lermet kapitain. Lannek Lermeti egin martxa-aginduan hauxe zehazten zuen : Laborde jabeak jakinarazi zuen nahiaren arabera, eta gerlarako armatzeko eman zituen aginduen arabera, helburua bazuten presa bat edo beste egitea, bidenabar. Ontzi etsai bat edireiten balin bazuten, kontuan hartu beharko zuten ingeles nazioarekin baizik ez zeudela gerlan, eta beraz nazio horretako ontziak baizik ez zituztela hartu beharko. Gogoan bazuten beraz harrapakina egitea, ahalaz.

Aipatzen den Jean Laborde, ontziaren jabea, goimailako pertsonaia zen. Bidarten sortua 1710an, salerosle handi, altxorzain, Erregearen prokuradore Louisbourgen. Berantago preso egonen zen Bordeleko Hâ gotorlekuan, dirua desbideratu zuelako akusaziopean; Eysines herrian hil zen 1781an. Aipatu dugun Martin Lermet kapitaina Ziburun sortua zen 1731an eta beraz gaztea 1757an. Aita ere kapitain izan zen Ternuan eta kortsario. Ordezko kapitaina Dominique Leblanc ez zen euskalduna, 1731an Louisbourgen sortua. Lotinanta Pierre Derratsou zen, 1732an sortua maniñel frangotako familian. Baionak orduan 2500/3000 mariñel en-

pleatzen zituen eta Donibane Lohizunek 800 bat. Hainbeste mariñel biltzea ez zen erraz: Baionako burgesek beren gizonak igortzen zituzten Hegoaldera bereziki mariñel berriak biltzeko. Baina *Le Dauphin*-en kasuan, eskifaia-zerrendak salatzen duenaren arabera, ontzian joan ziren mariñel gehienak bertakoak ziren, Lapurdikoak, eta haien patronimoak gaur ere erabiliak dira, hala nola Bizente Lisssaratsou, Donibane Lohizunekoa, Samson Hiriart, Azkaingoa, Bertrand Guichendeguy, Senperekoa. Ba zegoen ondoko herrialdetatik etortzen zenik, adibidez Ortzaizeko Jean Detchebere.

Le Dauphin ontziaren abiaburua Baiona izanik, goazen ikus zertan zen ontziak zeukan helburua, hots Louisbourg herria. 1713ko Utrechteko itunaren ondotik Frantzia behartua izan zen kolonia batzuen Ingelesei abondonatzera, hala nola Akadia eta Ternua. Baina 1757an Louisbourg oraino Atlantikako munduari idekitako leihoa zen frantsesentzat. Louisbourgko garrantzia ongi deskribatzen du X. Lamikiz (Lamikiz 2015). Arraintzan, bakalau gatzatuaren exportazian eta Europatik etorraraziko ekoizpenetan oinarritzen zen bertako ekonomia. Jendeen aldetik, gotorlekuak behar zituen soldadu asko gerizatu; arrantzaleen artean badakigu Euskaldunak ez zirela guti : 1752 ko errolda baten arabera 199 arrantzale bazeuden Îsle Royale delakoan, eta erdia osatzen zuten Euskaldunek; Saint-Malo aldetik jinak gutiago ziren %37 baitziren. Ofizialeen eta salerosleen artean %15 bat euskaldunak ziren.

X. Lamikizek hatzeman euskarazko gutuneria inguru horretan kontuan hartu behar da. Kontinente batetik ugarte batetara joatekoa zen gutuneria, ingeles ontzi armaturik ez balin bazen izan biderdian. Gutunak ez ziren heldu eta Londresen gorderik daude 221 dokumentu Le Dauphin ontzian zeudenak. Anpersana deitu Ikerreko webgunean kontsultagarri izanen dira osoki hementik laster. Erranikako 52 euskarazko gutunez gain, 108 gutun pribatu badaude, frantsesez eginak, eta 33 dokumentu administratibo: haietan, eskifai-errolda, Penthièvre almiranteak eman baimena etabar. Ondotik nahas-mahas beste dokumentu batzuk eta azkenik, auziari buruzko 9 pieza bildu dira, ingelesez idatziak.

19 mendeko euskarazko ekoizpen idatzia ezagutzen duena ohartzen da beste nonbait xume iduri lukeen 52 gutunetako korpusa ekarpen handia dela. «Seismo» eta «tsunami» hitzak erabili ditu Céline Mounolek (Mounole 2023) ekarpen honen pizua haztatzeko, zeren eta txiki eta labur izanik ere, korpus honek ez baitu parerik euskal munduan. Gainera aro zehatz baten hizkuntza-egoera bati buruzko informazioa ematen du (1757ko martxokoa). Garatzen duen lan linguistiko interesgarri batean, C. Mounolek erakutsiko du zein distantzia, zein tramu den gutun horietako hizkuntzaren eta orduko literaturako hizkuntzaren artean. Euskarazko literaturaren leku eta momentu indartsuak eta ahulak luzaz aipatu gabe, orroit gaitezen hastapeneko literatura orozbat Iparraldean gertatu zela, eta Lapurdin gehienbat, Etxepare kendurik. Hegoaldean Larramendi aipatzen da euskarazko ekoizpenaren mugarri gisa, 1729 gramatikarekin eta 1749ko hiztegiarekin. Hain zuzen ber denboran, *Le Dauphin* ontziaren murrizten hasten da Lapurdiko ekoizpena. Ahuldura horren azaltzeko Lafittek bereziki Utrechteto itunaren ondorioak aipatu zituen, honek krisi ekonomikoa ekarri zuelakoan. B. Oiharzabalek erakutsi du ez dela hori oinarrizko arrazoia eta diglozia kontuan hartu behar dela: euskarak, gaztelerak, frantsesak, latinak, okzitanerak ez zuten pisu bera gizarteko esfera desberdinetant. 17. mendean Materrek, Etxeberrik, Axularrek euskaraz idazten balinbazuten, helburu garbia bazeukaten: bazegoen orduan euskara baizik ez zekien irakurlego bat eta hari hazkurria eman behar zitzaion euskaraz. Behar hori bete zuten. Irakurlego hori nolakoa zen, zein jendek osatzen zuen ez dakigu zehatz. Baina dudarik gabe jadanik beste irakurlegoa bazegoen Chartier-ek «arropako jendea» deitzen duena ; goimailako apezak, nobleak, militarrak, lumazko ofizierak biltzen zituzkeena. Aldiz, euskaldun hutsen irakurlegoan sartzen bide ziren laborariak, ofizialeak, salerosleak, mariñelak, burgesen parte handi bat. *Le Dauphine*ko gutunerian ikusten da gutun-igorleak ez zirela baitezpada pobreak edo eta gizarteko mail apalenetakoak.

Gutunak nuntik nork igorri eta nork hartu behar zituen badakigu. 52 gutunetan, 11 Saratik ateratzen dira, 8 Getariatik, 8 Senperetik, 6 Donibane Lohi-

zunetik. Emaile txikiagoak dira Ziburu (5), Azkaine (5), Hendaia (3), Ahetze (2), Urruña (2), Bidarte (2), Baiona gutun bakar batekin (nahiz ez den Baiones batek egina). Ohart, eta ez da ustegabean gertatzen, herri horiek guziek osatzen dituztela orduko Lapurdiko Biltzarrak izendatu zituen itsasaldeko herrien zerrenda eta eremua.

Igorleen sexua kontuan hartuz, emazte igorle gehiago bada gizon igorle baino: 35 emazte eta 12 gizon. Aldiz, gutunen hartzaileetan, gizon gehiago badago emazte baino: 44 gizon eta 4 emazte. Kontsideratzio orokorrik egin gabe, lagin horrek baldintza bereziak jasaiten dituela azpimarratu behar da. Euskal Herritik joan jendeak gizonak dira gehienik, arrantzale edo mariñel edo ofiziale. De Lancre de Rostéguyren hitzak hartuz, itsasoa da «biderik gabeko bidea». Hauxe gehitzen diogu: itsasoa bidea izateaz gain, irabazbidea da lur-leihorrak ez duelarik nahiko emaiten. Bide eta irabazbide horietan, gizonak gehiago ziren ziren emazteak baino. Aldiz, Euskal Herrian egondakoak (gutunetako) emazte dira gehienik. Beraz, ondorioz, goiko zenbakietan gutun-igorleetako emazteen kopuru handia ez daiteke extrapolatu emazteen alfabetizazio-tasa handia zela argudiatuz. Ez dela hala ongi erakutsi du X. Elosegik. Berak ematen duen adibide baten arabera, Saran, 1758an, ikertu dituen 48 bataio-agirietan, 15 gozaita (aitatxi) agertzen dira firmatzen dakitenak eta aldiz goxama bakar batek firmatu du. Ezkontza-agirietan, 8 esposjaun badaude firmatzeko gai direnak eta emaztegaietan bihirik ere. Donibane-Lohizunen, kostaldean, alfabetizazioa ez da hain apala. 44 ezkontza-agirietan, 26 gizon badaude firmatzen dutenak eta 10 emazte. Beraz, emazteen alfabetizazio-tasa handiago da euskaldun gutiago den herrian, Donibanen, baina herri honetan ere, gizonen eta emazteen arteko diferentzia esanguratsua da. Hori kostaldean nola barnekaldean (hemen Sara barnekaldekoa dela kontsideratuz). Informazio gehiago ekartzen du X. Elosegik baina ondorio berdintsuak erakutsiz.

Igorlearen eta hartzailearen arteko loturaz bi hitz eman behar dugu. Ez da harritzekoa gehienik agertzen den lotura da askazgoa, familia, ahaidetasuna eta bereziki aitamen-semealaben artekoa balin bada. Gutunen erdia egiten du hartzailearen aitak edo amak. 7 esposandere badaude senarrari idazten diotenak. Gero arreba, anaia, alaba, osaba, koinata agertzen dira eta erretora, auzoa, adiskidea. Gutun-hartzailearen ofizioa emana da frankotan: *marinel*, *zurgin*, *maiasturu*, *zapatain*, *kapitain*, *mutil*, *sehi*.

Goazen gutunen edukiaren ikustera. Zein ote da gutun baten idazteko zergatia, oro har? Erranen dugu gutun bat firmatzailearen bakarrizketa dela nahiz ber denboran idazten duenak elkarrizketa bat hazten duen hartzailearekin. Bakarrizketa eta elkarrizketa batzen ditu gutun batek. Igorleak dituen xedeetan ondoko arrazoiak edireiten ohi dira: berriak ematea, zerbait kontatzea, erantzutea; debekatzea edo baimena ematea, abisatzea edo gaztigatzea, agintzea, jakinaraztea (Delmas). Erran gabe doa xede bat baino gehiago izan dezakeela gutun batek helburu berezi bat edo beste azpimarratuz. Doinuaren aldetik, omore tzarra erakusten dute batzuetan eta agindu gogorrak ematen dituzte igorleek hartzaileak izan zukeen portaerarengatik hala nola berririk ez ematea, laguntza ekonomikorik ez igortzea, errespeturik gabeko jokatzea, gaizki ikusitako amodioak «bestaldean» izatea, herrira itzultzeko gogorik ez erakustea. Doinua lasaiago dute gutunek familia zuzenki hunkitzen ez dituzten gertakarien aipatzeko hala nola Erregeren kontrako atentatua, kortsua eta gerla, herrian gertatu istripuak etabar. Aitortu behar da gutunek ez dutela deus ere kontatzen luzaz: edo eta epe laburra zutelako gutun igorleek beren idazkia emateko edo eta halako erreserba bazutelako kanpoko gertakari horietaz mintzatzeko eta are gehiago beren sabeleko mina aipatzeko. Hitzjarioa handia den ala aski murritza ez dugu sakondu. Iduritzen zaigu gutun luzeak bakan agertzen direla, alusioak maiz erabiltzen direla, lakonismoa handia dela. Erdarazko idatzitako gutunak ez ote diren luzeagoak barnatzeko da. Ikertzeko da ere eskribanoaren (letra bere eskuz idazten duena) eta gutuna diktatzen duenaren arteko harremana.

Gutun guziek gutieneko ekonomia badute: hastapenean agurtzeko formula bat etortzen da, garatua edo ez. Ageri da frantsesezko gutunetan ematen den agurtzeko formula badutela eredu. 20. gutuna (gutunen zenbakia ez da Lapurdum 2015 liburuan

agertzen dena: Anpersana wegunean agertuko dena da) hartzen balin badugu, paperaren aurkian idatzia da hartzailearen helbidea, frantsesez idatzia (gutunazalik ez da, gehienetan). Binperrezko partean dator testua, gorago aipatu dugun gutieneko *incipit* batekin, informazio laburra, eta agurtzeko formula laburra gutunaren bukatzeko, firma eman aitzin. Alabaina, Getariako Maria Lamark igorleak semeari idazten dio, haren berririk ez duela kexatzen da, jakinarazten dio *çhampelak* (bakalaua harrapatzeko aparailuak) ordaintzeko bi libera hartu dituela, beste xanpel batzuk igortzen dizkiola. Bitartekariak aipatzen ditu: Borda etxeko Ahex eta Zantzingorriko etxeko jauna. Firmatzen duelarik, alarguntsa dela aipatzen du. Horra testua:

> *Lapresente lettre soite / reandeue a Bernat de Lupoy çhentiqa / a Louisbourg de Louis[bourq] fait aguethary, cé 10 Marz 1757 : «Enne seme maitea gastiaxen dauxut milla gorainxy Estut Erreçibitu çure ganiq aurten Es Letrariq Es çolament gorainxibat jçatu ditut 2tt ahex Bordaganiq chanpeletaqo Eta Egorxen dausquixut çhanpelaq çançingorrieneqo jaunarequin. Gueldixen nais cure ama allargun gaçhoa çure çerbitçary mariade Lamarq.»*

Sarako Martin Borda igorleak Andreco Borda semeari egin gutunean (12.gutuna), incipitaren eredua osoago da. Darabilan estiloa aski ona da, gutun batetik bestera aski fidelki errepikatzen den eredua jarraikitzen duela nabari da.

> *«Ene seme maitea, hartzen dut libertate zuri bi lerroren eskribatzeko eta zuri gure familiako berri onen gaztigatzeko. Osasuna dugu Jainkoari esker, eta desiratzen ginuke denbora berean zuk ere hala bazindu.»*

Gutun batzuen egileak hartzailea erasiatzen du. Bereziki hartzaileak ez duelakoan aspaldi handian bere berririk eman, gutun igorlea egoera ekonomiko latz batean dagoelarik. Horra Hendaiatik Mari Etcheverryk igorria (3. gutuna).

> *«Eskribatzen dizut bi lerro hauk zuri markatzerat emateko estatu miserable batian kausitzen naiz. Errezibitu dut zure letra ere; han ikusi dut zure berrijak, baiñan ez nai bezala. Martinek eta hunzija galdu dute berriz ere, horra zu, desira du arekin zerbait izanen duzu. Marikatania asarre da zure kontra, eztiozula letra bat egin lau hurte huntan. Alaba bat baduzu iduri dubena perfek; etorriko da horrat bakija egiten denian, idurikitzen du. Emendik harat oroitu zaren baiño obeki oroitu zu zarela gizona […] Gana Xume asarre da zure kontra, merisi du zureganik; gusama ere bai».*

Erasiak agertzen dira hartzailea gaizki portatu delakoan. Horra 43. gutunean, Urruñako Haraneder apezak Pierre Etchegoyen ilobari igorria:

> *«Uste nuben aditu denbora guzietan zutaz gizon behar den bezalako bat hizaren ziñela. Bizkitartean kontrarioa aditzen dut; ainitz gare doloratuak zutazko gauzen aditziaz. Ez dut dudatzen konpania gaiztoen efetu batzuek direla horiek; zorigaitzez hizan zarela herri hortan. Bainan zer egin presentian? Erreparatzerat entsaiatu behar da iraganaren, oraino gaztea zare. Hortarakotz, ene haur maitea, moienik hoberena zait alde huntarat pasatzea. Ethortzen bazare borondate on batez betea, entsaiatuko gare gauzen erremediatzerat, eta guk hizaren dugu zure alderako bihotz ezagutu daukuzuna. Ez dezazula, beraz, falta ethortzetik zure anaiarekin».*

Batzuetan hartzailea Louisbourgen ezkontzekotan dagoela jakinik, arranguratuak eta kezkatuak dira aitamak. Berri horrek esondatzen du hartzailea Ameriketan jarriko dela betikotz, etxekoenganik urrun. Mollères andereak idazten die Bitor eta Bernat de Farandiret bi gizon horien arreba batek beren ezkontza prestatu duela herrian erranez (7. gutuna) :

> *«Zuen arrebak egin du hekien ezkontza; zuenak ere egina dago, baldin disposatzen bazarete. Bitor, ene anaia maitia, konpreniarazi izatu daut marinel batek hor habitatzen zarela, bainan guziarekin ere behin ere eztut sinetsi, zergatik ezpaitzaozkat kapable zu guri damu horren egiteko. Beinat, ene bihotz maitia, dio zure arrebak baduela hirur urte zure emaztegaia bilatua duela, bethi dagoela hagorandua zu ikusi nahiz, bethi galdegiten diozkala zure berriak eta bethi agintzen zaituela, baldin fraide egin gogo ezpaduzu. Ezkontza tratazale handi bat da hura, bainan berrerik eztu egiten, hargatik zuek egiteko dauka hura».* Habitatu hitzaren erranahia ez da «bizi izan». Erran nahi du «konkubino batekin bizi» eta hortik dator igorlearen hasarrea.

Pentsatzekoa den bezala, erlijioa agertzen da eta kristau moraletik ez baztertzea gomendatzen diote igorleek hartzaileari. Horra Joanna Castillouk dioena, hartzaile den aita Jainkoaren borondatearen arabera plegatu behar dela gomendatzen baitio :

> *«Ene aita maitia, bihotzez zuri errateko hemengo dolore eta miseria guziek baino gehiago aflijitzen naute zuhor gerla triste hunekin izanak eta zu ikusi nahiak; zure berriak ezin jakinez hiltzerat goaz guziak. Ene aita maitia, egizu bethi kuraie; zeruko Aitak bethi izaren du gure artha; eta zure dolore guziak ofreji diotzotzu Jinkoari ahalik hobekiena. Zure horko doloriek ematen dute ene bihotzian penarik handiena. Ageri da Jinkoak horla plazer dubela; konpli bedi, bada, bethi Jaunaren borondatia.»*

Erlijioa eta ohar mistikoak utzirik, beste gai bat errepikatzen da gutunetan: hartzaileak laguntza ekonomikoa igortzea, are gehiago igorlea nekez bizi delarik ahal handirik gabe. Horra Saratik Joanna de Bidartek Joannes d'Etchetori eskatzen dioena (23.gut.)

> *«Doloratuak gara hagiz izatu dugun berri tristeaz. Harren, gauzak ahalik ungi altxaturik, hunat pasatzerat egizu lehenbiziko komoitatean, zeren nihor gabe kausitzen baikara. Katalin eta bi-biok gara, eta Maria eta Arrangoitzen gelditu dire. Hunat pasatzerat egizu, zeren heldu den neguko ez bazaitzaigu ethortzen utz ahal baitezakegu errota.».*

Gutunaren hastapenean aipatzen den berri tristea zein ote zen argitzeko Xabier Elosegi ibili da Pirinio Atlantiketako arxiboetan, gutunetan aipatu protagonista askoren arrastoa bilduz. Gutun honen kasuan, dokumentu erabakiorra hatzeman du, hain zuzen 1756ko abenduaren 24an Sarako elizan ospatu baitzen zeremonia bat Pierre Etcheto zenaren alde: «fils des meuniers du moulin appellé d'Olha, noyé à Niganich en terre neuve suivant la déclaration de Joannes Etcheto frère du défunt dans sa lettre du 20ème novembre dernier écrite de niganich et adressée à sa mère au moulin d'Olha ». Pierre Etcheto Gaspesiara joan zen 1752an arrantzale gisa. Artxiboetako dokumentuak aipatzen duen Olhako errota da gutunean agertzen den errota bera: Olha errotatik ez ote den bortxaz joan beharko beldur dela dio gutun-igorleak.

Beharrik gertatzen zen igorlearenbaitara laguntza etortzea. Gehienetan kanbio-gutunen bitartez. Testuetan *letra-sanja* deitua da. Catalin Lalannek horrela idatzen dio Saratik Esteben Gourdo senarrari : « *Letra hau eskiribatzen darotzut zuri erraiteko errezebitu nuela zure letra 120 liberako letra sanjarekin, zeintaz pagatua izan bainaiz. Bertze alde, erranarazi tut zuk markatzen zinarozkidan mezak* (6. gutuna) ». Senarra 1708an sortua, han nonbait, Bresten hil sen 1758an Gaspesiatik itzultzean, X. Elosegik erakutsi bezala. Bordariak ziren Lekuberria deitu etxean eta Sarako Etxeberri mediku eta idazle famatua etxearen jabea zen. Beste kasu batzuetan dirua izateko trabak badaude. Horra zer dioten Brave aitamek Senperetik: « *Horko letra sanjarik Baionan ez dute errezibitu nahikatu. Arrotxelarat igorri omen tuzte, ez dakigu han errezibituco othe tuzten* (10. gutuna)". Eta dirua eman behar zukeen arduradunaren fede ona dudan ezartzen dute ere hala nola 11. gutunean. Hendaiatik, Gratianne Diharcek kondatzen dio Joannes Hiraboure senarrari nola egon den Biarritzen hartu gogo zuen diruaren bila eta nola delako Manesca jaunak ez duen onartu bere funtsa ematea:

> *«Izatu dut zuk egorri dautazun letra xanjia, non markatzen baititu berrogoi luis, zeiñetan adrezatuba baita M^r^ Manesca Miarritzkuarenganat. Beraz, izatu naiz jaun haren baitan, nola egorri dautazun, eta errepusta eman darot nola eztuben xanje egortzaille horrenganik deus funtsik errezibitu, untzijak galdu direla, batzubek eta bertziak ez ageri non diren. Ez darot sinatu, ortarik hartuko ditutzu zure prekozijoneak. Ene espos maitea, egiñen duzu hahalik hobekijena, ezen iduritzen zait pagatzeko borondaterik izatu balu sinatuko zubela».*

Beste zentzuan ere eman daitezke laguntza, Louisbourgen biltzen ahal ez diren gauza premiazkoak Lapurditik igorriz: arropak, janariak, edariak, arrantzarako tresnak, haziak… Martin Durrutyk Domingo semeari horrela erraten dio (15. gutunean):

> *«Eta igortzen darozkitzut urdaiazpi bat eta lukhainka zenbeit, eta andoilla bat». Edo 30. gutunean: «Anaia, egortzen dautzute aitak eta amak Picunen semiarekin zerbait aldatzeko: galtza, maripolisa, athorra, marra, galtzerdi eta zapata. Eta ene partez emanen dautzu Martin Filisen semeak kintal erdiko barril bat*

nafar, eta Tulikunek emanen dauzkitzu barril xume bat agorient eta urdaiazpi bat, ene esposaren partez».

Ohartarazten dugu gaur enblematikoak diren Nafarroako ardoa eta Baionako azpia jadanik aipatuak direla igortzen diren opari horietan.

Beste berririk ere ematen da gutunetan, bereziki kortsuarengatik familiako edo auzoko norbaitek bere larrutik ordaintzen duelarik, hala nola preso egotea. 22 gutunean, *Ganis on* gizonaren zortea aipatzen da.

«Ene haur maitea, Ganis ona gan tzunan kortsura, eta angelesek harturik etxatu ziztenen Espanijarat, Lisibona erraten dioten portu batera, zenak gasoa etorri baitzen onez berrehun lekoatan bere lagunekin pilala triste batean. Etzunan hura ikustea baizik bertze mirakulurik, buluzgorrija hurren!»

(Bide nabar, hizkuntza dela eta, alabari idazten duen amak nokako aditz-tratamendu alokutiboa erabiltzen duela ikusten da).

Aldiz kortsuak idekitzen dituen diru egiteko bideak aipatzen dute beste gutun batzuek. Kortsua lanbide ona dela eta mozkin ederra ekartzen duela segurtatzen du Pierre d'Etcheversek (30. gut.) bere anaiari gomendatzen dioelarik herrira itzultzea negozio ona egiteko ontzigintzan:

«Hanitz dire trabailu handiak Frantzian, untzi egiten, partikularki Baionan, eta Donibanen ere bai, ezen Donibanen egin ditut nik neronek korsuzaleak nere planarekin, bat Baionako eta bertzerik hirur edo lau aski untzi galantak, Erregeren konstrikturaren planaren ganean. Eta orai ere baditut hirur xume xantierean hasiak, eta handi bat badut hasi beharra ene kontuko egiteko, zeinak 18 peza muntatuko baikitu».

Ikusten da gutun-igorleak aski ontsa bizi dela gerlari esker eta ez ezin dela familia hura kokatu gizartearen pobreen artean.

Lekuko gertakariak ere aipu dira, aski zalu. Nor sortu den, nor hil den. 6. gutunean auzo bat hil dela agertzen da: "*Ortholopitzko andrea Jinkoak deithu du bereganat hil hunen hirurean, San Blaz egunean, andre gaztea*". Gaur ere, Ortillopitz deitzen den etxe hori ezaguna da Saran museo ttipi bat balitz bezala bisitatzen baita.

X. Elosegik ikerketa sakona egin du bi gutunetan agertu gertakariei buruz (Elosegi, 2019, 92-113). 45. gutunari esker, norbaitek Sabat Dorre deitu Ahetzeko gizona hil duela jakiten dugu:

«Eta zuri hemengo berrien adiarazteko, segur da aditu izan tuzula hemengo berri tristiak. Ehaile gaixoa preso harturik Jondone Laurendi egunian, 1755eko urthean, Parisat eraman dute, Sabat dorreko zena norbaitek hiltzea dela medio, eta bi hilabete badu jendek darazatela bidean heldu dela libro bere etxerat, baiñan oraino ezta ageri gaixoa. Eta Ostaleriako premua, urthe berian Jondoni Jauni inguruan preso harturik, Toulonen da kondenaturik galeretarat sekulakotz, adiskideak ongi melaturik ez urkatzeko. Eta diote denbora batez harekin ibili zen gizon bat, orobat, hura bezala jujamendu beraz kondenatua dela, Amizola ohiaren kolpatzea eta Seroraren leiho hausteko sujeta dela medio».

Ehaile «*gaixoa*» dio gutunaren egileak, ehaile hartaz urriki balitz bezala. X. Elosegik aferaren testintungurua eta harat hunatak deskubritu ditu: falta zaio bakarrik hiltzailearen izena zehaztea. Parada baliatuz, AD 64ko artxiboetan X. Elosegik bildu ditu euskarazko dokumentu interesgarriak, afera honi buruzkoak eta besterik ere, ezagutzen ez zirenak.

Bigarren afera batean (25. gutunean) ez da ehailerik baina bai emazte baten hiru senar bata bestearen ondotik zendu direnak. Catalin Marie Berrogain deitzen den emazte baten alaba hiru ezkontzen bidez eta dotea gero ta pollitagoaren jabetuz igoera soziala nola erdiesten duen erakusten ditu X. Elosegik. Alaba hori, Marie Etxeto, lehenik 1741an ezkondu zen Baionako salerosle zen Pierre Saint-Genesekin. Doteak 12 000 libera ekarri zizkion. Senarra hil zen 1745an 72 urte zituelarik. Alarguntsa berriz ezkondu zen 1747an Pierre Dufau ontzikapitainarekin. Doteak 25 000 libera ekarri zion eta emaztegaiak horrenbeste. Dufau senarra ondoko urtean hil zen, 1748an. 1751an hirugarren aldiz ezkondu zen Marie Detcheto Pierre Laborde d'Arbrun noblearekin. Hau Landesetakoa zen (Saint-Sever). Emazte horren traiektoria ezagutzen dugu Elosegiri esker baina *Le Dauphine*-ko gutunek akulatu gaituzte orduko gizartean murgiltzera. Erran bezala, personaiak ez dira baitezpada ahal ekonomikorik gabekoak: 45. gutun hori idazten duen Marie

Berrogainek Catherine Laborde andereari igortzen dio eta Catherine Laborde hartzailea Jean Laborde Louisbourgko kargudun handiaren arreba da.

Ordu da gutunetan behin baino gehiagotan agertzen den gaiaz hitz bat eman dezagun. Badago gutun horietan guzietan leitmotiv bat patetikoa. Ezen Euskal Herrian egon direnek hanitzetan hauxe diote Ameriketara joan denari: Zato! Zato! Zato ! Erretzen dira «hagorantzen» dira lagunak, ea noiz itzuliko den aita, anaia, arreba, auzoa. Bederen gerla bukatuko delarik etor baledi han dagoena, gaitzerdi! 7. gutunean:

> *«Martxoaren lehenbiziko igandian egin behar tu bortz urte hemendik partitu zinetela. Geroztik ezkara hala satifatu, nola aurten Jainkoak grazia handiak egiten darozkitzue, eta eztarotzue batere gutiago eginen baldin ematen badiozue satifazionea zuen aitari eta amari zuen ikustekoa, ezen hurren dire hagorandu zuek begiz ikusi nahiz: munduko ontasun guziak baino nahiago lukete zuek begiz ikusi, gero hilen badire ere. Hain ungi zuen arrebak zuen ikusteko desira baluke, non uste balu munduko tresorik handienak, guziak utz bailezake zuen ikusteko, non Bordele bezain hurbil balitz ez bailuke batere herabe urratsa zuen ikusteko».*

49. gutunean ere, Graciannek dio:

> *«Amak eta kuñatak milla gorantzi, eta koñatak othoi zatozela etxerat, ethortzerat posible bada niholere, hemen ere badela inobres f...go […], nahi dubenak khortsurat eta nahi dubenak martxantean 60 libera ilabetean, eta khortsurat hilabetentza 150 libera; nahi dutena egiten dute. Zato ahal badagizu etxerat. Zure ama maitea lehenik eta gero gu hurren, zuk aditzerat eman darokigutzun berriak aditu eta, ondoban zu horko miserietarik etorria ikusi nahiz etxerat; zato etxerat. Bi seme izatu tugu; lehenbizkoa hil tzen sortu ta zortzigarren egunean, eta bertzea haur galant bat dugu. Elkarri eginen diogu begitarte. Zato etxerat».*

Nahiz datu zehatzik ez dugun, badakigu gutunigorleek gutiziatzen zuten itzulera ez dela askotan gertatu. Gainera, ez dezagun ahantz gutun horiek hartzailegaiek ez dituztela sekulan irakurri ahal izan. X. Lamikizek ideki zituen gutun horiek beste nehork ere 250 urtez ez zituela ezagutu dakigularik, aitortu behar da halako zimiko bat ematen duela.

Azkenik, *Le D*auphin ontziko dokumentazioak beharleku interesgarri eskaitzen du. Orain artio egin ditugun ikerketek altxor hori ez dute osoki ustiatu. Hizkuntzalaritzaren eta filologiaren aldetik gutunen hizkuntza sakondu behar da usaiako alderdietan: ortografia, lexikoa, morfosintaxia, estiloa, pragmatikoa etabar. Bestalde zehaztu beharko da zein den orduko literatura idatzian erabiltzen den euskara «kultoa»-ren eta hemengo gutunerian agertzen den jende «soilen» euskararen artean dagoen distantzia. Ba ote da orduko aro diglosikoan alfabetizazioa eta eskolatzea neurtzeko modurik? Frantses erresuma kontuan harturik, zergatik bertako hizkuntzan egindako gutuneria erakusten ote du Lapurdi bezalako lurralde murritz batek, halakorik ez baita agertzen Bretainian, Okzitanian, Katalunian —orai ditugun datuen arabera, behintzat—?. Eta zergatik hegoaldeko probintzietan ez da euskarazko gutun asko bildu —dagoenekotz— Londreseko artxiboetan? Beste mail batean erran behar da X. Lamkizen lana segitzeko dela Atlantikoaren garrantziari buruz, komertzioari eta kortsuari buruz X. Elosegik nagusiki ideki duen ildoak behar du barnatua izan, Elosegi beraren eta ikerlari berrien eskutik. Itsasoaren bi alderdietan eta gutunetan aipatuak diren pertsonaien errolda aberastuko da orduko gizartean hobeki ezagutzeko, nolazbait intimitatea apur bat hunkituz.. Bestalde X. Lamikizek, A. Lafuentek eta Ikerreko proikektuaren arduradunok jarraitu behar dugu Londreseko TNAko dokumentazioa aztertzen, bere osotasunean, hots beste hizkuntzetan eta beste eremuetan dagoen altxorra kontuan hartuz. Euskarak asko badauka irabazteko Euskal Herriaz kanpoko ikerketekin lotura hertsian egonez. R. Cock TNAko arduradunak legunki erraten duen bezala, jakitate-arlo horri lotu zaizkion kontinenteko ikerlariak ez dira hain ugari izan. Ordu da gure taldeak erronka hori har dezan helburu eta xede.

Bibliografia

Carette, Camille, 2019 : *Editer des sources multilingues du XVIII siècle en XML-TEI : le projet le Dauphin*, Mémoire pour le diplôme de Master Technologies numériques appliquées à l'Histoire, Ecole Nationale des Chartes, 96 orr.

Chartier, Roger, 2015 : « Entretien avec… Propos recueillis par Évelyne Cohen, Pascale Goetschel », Sociétés et représentations, 40, 289-321.

Cock, Randolph, 2023 : « La capture et le procès du Dauphin », (argitaratzeko) 2023, à paraître

Cock, Randolph, 2023 : « Les prize papers », (argitaratzeko).

Delmas, Bruno, 1997 : « Correspondre : esquisse d'une typologie des formes individuelles et collectives de la communication écrite » in Pierre Albert (arg.), *Correspondre jadis et naguère*, 120. ongresua, CHTS, Aix-en-Provence, 13-30.

Denis, Gwendal, 2023: « *Ar Baum Glas,* texte breton de 1757 », (argitaratzeko).

Elosegi, Xabier, 2019 : *Le Dauphin : 1757 ko gutuneriari buruzko osagarriak eta gogoetak, « othoi çato etchera,* Bilbao, Euskaltzaindia ; Iker bilduma.

Lafuente Annabelle Lafuente, 2021 : De Bayonne à Louisbourg : la correspondance du bateau bayonnais Le Dauphin durant la guerre de Sept Ans (1757), Master, UPPA (zuzend. Ph. Chareyre).

Lamikiz, Xabier, 2010: *Trade ant trust in the eighteenth-century Atlantic world: Spanish merchants and their overseas networks*, Woodbrigde, Boydell Press / Royal History Society.

Lamikiz, Xabier, 2015: «Le Dauphin itsaontziaren testuinguru historikoa: Louisbourgeko euskaldunak, ekonomia atlantiarra eta gerra -1713-1758)», in Lamikiz X. allii, *Lapurdum*, 17-43.

Lamikiz, Xabier, Padilla-Moyano, Manuel ; Videgain, Charles ; 2015 : « Othoi çato etchera : Le Dauphin itsasontziko euskarazko gutunak (1757). *Lapurdum*, ale bereziak 2.

Mounole, Céline, 2023: Les lettres du navire Le Dauphin: un séisme pour l'histoire de l'écrit basque (argiratatzekoa).

Oyharçabal, Bernard, 2001 : « Statut et évolution des lettres basques durant les XVIIème et XVIIIème siècles », *Lapurdum*, 6, 219-287.

Peytaví Deixona, Joan, 2023 : « Une correspondance personnelle en catalan dans la guerre de Sept Ans : la vie quotidienne de Roussionnais expliquée par-delà l'Atlantique (argitatzekoa) ».

Pontet, Josette, 2006 : Les corsaires dans l'Amirauté de Bayonne au XVIIIème siècle, *Revista de Estudios marítimos del Pais Vasco*, 5, 309-324.

Talec, Jean-Philippe, 2023: « Le jeu de données «Le Dauphin de Louisbourg (LDL)» déposé dans l'entrepôt Anpersana (argitaratzekoa) ».

Viaut, Alain, 2023 : » La présence de l'occitan dans les fonds des prises de guerre au XVIIIème siècle des Archives de la Haute-Cour de l' Amiraute Britannique : une première approche (argitaratzekoa) ».

Diseinu Partizipatibo Erradikala: parte hartzeko prozesuen erdigunean komunitatea jartzen duen metodologia eraginkorra

Gizartean egituratzen diren parte hartzeko ohiko prozesuetatik harago doan metodologia bat da Diseinu Partizipatibo Erradikala (DPR), erdigunean pertsona jarri beharrean komunitatea bera jartzen duena. Prozesu luzeak izaten dira halakoak, baina frogatu da gai direla gizartearen krisiei erantzun berritzaileak eta integralak emateko. Arrakasta bermatzeko beharrezkoak diren gakoak dakartza ikerketak: botereari uko egitea, erakundeen erresistentziari erreparatzea, pribilegioei begiratzea eta zapalketek eragindako traumak zaintzea.

Artikulu zientifiko honek DPR metodologia berritzailea aztertu du, eta egungo krisi sozialei aurre egiteko baliatzea proposatzen du, izan ingurumen-krisia, pobrezia-zikloak eta krisi ekonomikoak zein gatazka sozial askotarikoak (arrazismoa, etnozentrismoa...). Izan ere, krisi horiei aurre egiteko, ez dira nahikoa norbanakoek modu indibidualean egin ditzaketen aldaketak, komunitate mailakoak izan behar dute. Beraz, erabakiak komunitate mailan hartzeko prozesuak bideratzea eta ekintza kolektiboak planteatzea du helburu DPRak.

Ohiko metodologia partizipatiboen aldean, norbanakoen partaidetza erabat erradikalizatzea proposatzen du DPRak. Inklusiboa izateaz harago, non jatorri, adin, esperientzia eta ikuspegien dibertsitatea bilatzen baita, komunitateko norbanakoak prozesuko kide oso bihurtzen ditu, diseinatzaileen maila berean kokatuta, ahots eta erabakimen-maila berarekin. Ondorioz, diseinuaren fase guztietan aktiboki hartzen dute parte, hasi definizio-fasetik eta inplementazio-faseraino. Egindako lana berdin ordaintzen zaie komunitateko kideei eta diseinatzaileei.

Diseinatzaileek botereari uko egitea da prozesuaren gakoetako bat. Ez badiete benetako ahalmena ematen komunitateko kideei edozer aldatu eta erabakitzeko, ez da DPR prozesu bat izango. Gizartean boterea erakunde eta pertsona jakin batzuetan egituratzen denez, ezinbestekoa izaten da etengabe egitea botereari uko egiteko ariketa kontzientea, bai erakundeek, bai diseinatzaileek eta baita norbanako pribilegiodunek ere. Boterea duenak ez du besterik gabe askatzen boterea.

Pribilegioak eta zapalduen traumak

Prozesu horietan, ezinbestekoa da pribilegioen kontzientzia izatea. Pribilegioan eta desabantaila-egoeran dauden pertsonak elkartu eta botere-maila berean kokatzen direnean sortzen da komunitatearen potentzialik handiena. Alabaina, kontuan hartu behar da komunitate kolonizatu, zapaldu eta baztertuetako kideek bizipen traumatikoak izaten dituztela askotan; beraz, prozesua bideratzen duten diseinatzaileek gai izan behar dute trauma-seinale horiek identifikatzeko eta zaintzeko, kide horiek ahalduntzeko.

Bestetik, ikerketak ondorioztatu du garrantzitsua dela prozesuei behar adina denbora ematea. Komunitatearen erritmoa errespetatu behar da, eta funtsezkoa da pazientzia eta malgutasuna izatea. Parte hartzeko ohiko prozesuak baino luzeagoak izaten dira honelakoak, normalean.

Baldintza horiek guztiak bermatuz gero, komunitateko kideek beren jakinduria eta kezkak mahaigaineratzen dituzte, eta prozesutik ateratzen diren irtenbideak integralagoak eta jasangarriagoak izaten dira. Beraz, oso baliagarriak izan daitezke egun gizarteak dituen krisialdi sozial, ekonomiko eta klimatikoei erantzun bat emateko.

InpAkta 00 (2023) 169-187

https://doi.org/10.47061/jasc.v2i2.3816

Radical participatory design: awareness of participation

Victor Udoewa

Chief Experience Officer and Service Design Lead in the NASA Small Business Innovation Research and Small Business Technology Transfer (SBIR/STTR) Program

1. Introduction

As a human animal, a part of nature, I inhabit multiple spaces of privilege and lack of privilege. I am a cisgender, heterosexual, Christian, male, U.S. American human. Simultaneously, I am a Black, disabled Nigerian in the U.S. from an immigrant family. I am a member of the indigenous Ibibio people group, and my name, *Anietie*, is a shortened version of the phrase "Who is like God?" When I write, I tend to write from a perspective of African indigeneity, different from indigenous perspectives in the Americas or Australia. There are many other parts of my background that place me in positions of privilege or disadvantage-country of residence, education, income, etc. Many of those have changed throughout my life.

One influential privilege I hold is the position of designer. I have practiced design in communities around the world. Despite my highest hopes, design has not risen to the challenge of resolving our current, growing crises. We face an economic crisis locking some people, groups, and nations in cycles of poverty with fewer people controlling greater shares of the wealth; a climate and environmental crisis of ever worsening ecocidal devastation; a conflict crisis where entrenched casteism, xenophobia, jingoism, and ethnocentrism fuel ongoing disputes; and a spiritual crisis where none of our best faith traditions have been able to address any of the previous three crises. This crisis-bound world is a world of our monohumanistic design, creating a one-world world, in the service of industry and capitalism (Escobar, 2018; Law, 2015; Wynter, 2003; Wynter & McKittrick, 2015).

Awareness-based system change agents recognize that we cannot solve our crises only with external methods, or methods focused on creating change outside of ourselves. We must pay attention and nurture our inner life and interiority that provide the source conditions fueling our actions (Scharmer, 2009). However, it is not enough to do this on an individual level. It is difficult for social change to happen if I only nurture my interiority and no one else does, or if we each do it individually. We must also pay attention and nurture our communal in-

* **Corresponding author / Harremanetan jartzeko:** Victor Udoewa. CTO, CXO, and Service Design Lead in the NASA Small Business Innovation Research and Small Business Technology Transfer (SBIR/STTR) Programs. e-mail: victor.udoewa@nasa.gov. **Contributed by** Antonio Casado **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Udoewa, Victor. «Radical Participatory Design: Awareness of Participation», *Journal of Awareness-Based Systems Change*, 2022, vol. 2 n.º 22. (https://doi.org/10.47061/jasc.v2i2.3816)

teriority, directing our collective decision-making and actions. Participatory design (PD) attempts to focus on communal interiority. Instead of focusing on the external methods, methodologies, or what we do, it focuses on how we do what we do, the interior dynamics, ecology, and positionality of a living community.

If design has failed at creating a pluralistic, flourishing world, PD has experienced a type of stillbirth, never truly beginning to bring about the emancipatory democracy promised as one of its goals (Geppert & Forlano, 2022), struggling to rid itself from its inherent coloniality (Cooke & Kothari, 2001). This makes sense as our namologies —studies, types, or ways of designing— are simply a reflection of our ways of knowing which are a reflection of our ways of being (Ibibio, Generations). Our ontologies are our epistemologies and our onto-epistemologies are our namologies. Thus, there is a need to decolonize decolonization, or more specifically, decolonize PD. By embodying an onto-epistemic framework of relationality, the design process can become radically participatory. To embody relationality, designers need an awareness of participation, and, from awareness, can take action.

The first purpose of this paper is to go beyond critique, to decolonize and refuture PD. Secondly, I aim to holistically describe the PD I have experienced, as many PD researchers and writers do not often explain fully how, when, and what PD was implemented on a project. Third, through a holistic description, I want to place the PD I have experienced in comparison and conversation with what others mean or practice when they use the term PD. Lastly, I hope to encourage participatory designers to go further, fully radicalizing participation while encouraging non-participatory designers to begin the PD journey with a radical approach or goal. Communities, the people for whom professional designers design or the people who will use what is being designed, can and have always practiced radical versions of PD without professional designers. The problem occurs with the colonizing presence of professional designers. This paper presents not just insights but expertise from community practice that is not synthesized through mainstream, academic, institutional knowledge-based understandings of research rigor, but through indigenous, practical synthesis which is incarnationally and relationally codified by traditionalizing certain community practices and discarding other community practice, utilizing learning circles, storytelling, oral histories, art, ceremony, and more (Ellison, 2014; Smith, 2021; Suaalii-Sauni & Fulu-Aiolupotea, 2014).

In the rest of the paper, I share an awareness of the typology of participation and a description of Radical Participatory Design (RPD), the participatory meta-methodology this paper describes. Different from a methodology which is a collection of methods or guiding philosophies or principles that help one to choose a method at a particular point in a process, a meta-methodology is a way of doing a methodology, an approach or orientation that can be used with any methodology. Because RPD teams tend to gravitate towards certain methodologies over others, it can be considered an approach, orientation, or philosophy that guides one in choosing a particular methodology.

After introducing RPD, I discuss the ethics of RPD focusing on remuneration through the lens of equality and equity, dissemination of knowledge, and community accountability. I then discuss the evaluation of RPD to determine if the process was truly and critically radically participatory. This is helpful due to the invisibility of coloniality that may lead us to believe we, design team members, are practicing RPD when we are practicing colonial participatory design (CPD), conventional participatory design in which designers lead the process and participation is not fully through the design process. Next, I discuss the benefits of RPD, how it opens up pathways to other types of design such as society-centered, futures, systems, or planet-centered design. I share how it relates to empathy, comparing it to other design awareness-based systems change practices. I then discuss the difficulties of practicing RPD and provide tips to minimize the difficulties based on experiential knowledge. Lastly, I provide insights on addressing and overcoming organizational barriers to the practice of RPD.

2. The Awareness & Typology of Participation

To compare various PD practices, I use a typology of participation based on three spectra or questions (Figure 1). Who initiates? Who participates? Who leads? There is a temporal distinction between initiation, on one hand, and participation and leadership, on the other hand. Even though initiation only occurs at the beginning of a project while participation and leadership occur throughout, the effects of initiation can be experienced throughout the project, and initiation can even affect participation and leadership.

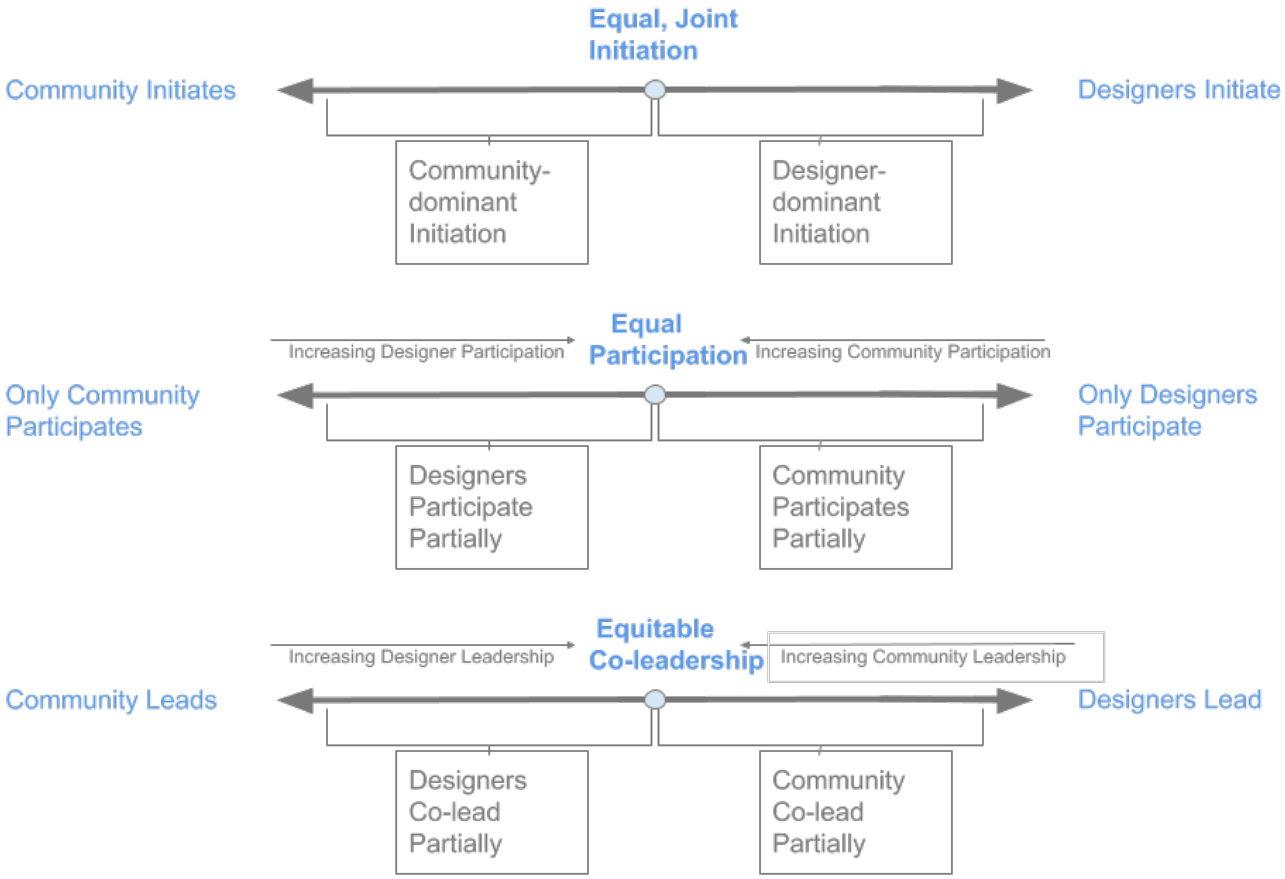


Figure 1
Three axes of participation: initiation, participation, and leadership

On the spectra, I locate: community design when only the community is involved; community-driven design when the community may invite professional designers for at least a little help, up to equal participation; CPD where designers fully lead and participate, never reaching equitable leadership with the community; and RPD in which the community fully or equally participates and fully or equitably leads. Visualizations for each type of design can be viewed elsewhere (Udoewa, 2022b in press). Radical comes from the Latin word "radix" meaning root. Radical Participatory Design is a design that is participatory to the root, all the way down, from top to bottom, beginning to end.

Thus, I introduce Radical Participatory Design as having three defining characteristics.

1. Community members are full, equal members of the research and design team from the beginning of the project to the end. There are no design team meetings, communications, and planning apart from community members. They are always there at every step and between

steps because they are full and equal design team members.

Communities are not homogenous. In RPD, we, the design team members, form qualitatively representative samples of the community in a way that honors cultural understandings of leadership and participation. We also drop designer-dominated notions of time, and move at the pace of community relationships, availability, and desire.

2. Community members outnumber non-community, professional designers on the design team.

   When a person is both a community member and designer, and she leads the process, choosing methodologies, she is practicing CPD. In RPD, when a person is both a designer and community member, she primarily embodies the community member role, offering design skills alongside all other community skills, while the community facilitates and leads the process. Because an organization may refuse to implement community ideas or prototypes during a PD project, there is a third characteristic.
3. Community members retain and maintain accountability, leadership, and ownership of design outcomes and narratives about the design artifacts and work.

   Characteristic 2 is a guideline, not a requirement. However, RPD projects tend to be more successful when they embody that characteristic. The goal of RPD is transformational justice, though RPD retains the benefits of multidirectional learning, inclusion of community perspectives, better design outcomes, and increased ownership over the outcomes.

In RPD, professional designers do not empower since empowering reinforces the hierarchy participatory designers seek to subvert with PD. Instead, in RPD, professional designers divest of power, and the community assumes it. The RPD process naturally becomes an educational one in which learning is embedded in every phase and activity, not just in research phases, due to the experiential, cultural, and spiritual knowledge the community embodies and their presence at every step in the process. Unlike research justice which views experiential, mainstream institutional, and cultural/spiritual knowledge as equal, RPD views experiential, cultural, spiritual, and embodied knowledge as greater or more important than mainstream institutional knowledge for system change.

Instead of the "designer as facilitator" model, we, RPD team members, move to a model of "community member as facilitator," "designer as community member," "community member as designer." "Community member as designer" means they are full-members of the team, researching and designing. "Community member as facilitator" recognizes that no matter how much designers try to neutralize our facilitation work, facilitation is power, and the power should be wielded and held by the community on whose behalf we are designing. So community members facilitate the process. Lastly, "designer as community member" signifies that the designer sits equal to and alongside all the other community members on the team, offering her skills (design and research) as equal to and alongside all other skills, assets, talents, and gifts of all other community members.

Through these models, RPD creates suspended space with an alternate social field. A social field is the structure of the social relationships between individuals, groups, organizations, and systems (Scharmer, 2009). Suspended space is a space where the social rules, norms, and relationships, governed by the larger society, are suspended in the subset space or small-group space within the society (Rollins, 2006). Because those social norms, rules, and relationships are different, an alternate social (sub) field emerges. Radical Participatory Design creates an alternate social field which aims to move across 3 stages. In the first stage, intrapenetration, the colonial logics of the macro-social field of the societal system naturally enter into the micro-social field of the design process. In the second stage, interpenetration, the micro-social field of the design process also begins to affect the macro-social field of the system and some of the new relationships held or suspended in the design process begin to carry over

into societal interactions outside the design process. In the last stage, extrapenetration, only the design process's social field is affecting the larger system's social field. The more RPD is practiced the longer the design team is able to sustain the suspended space naturally outside the design process or in other projects.

There is still more work to do to decolonize awareness-based systems change methods which are not yet or not necessarily radically participatory. The MAPA innovation lab (Sbardelini *et al.*, 2022), social field action research (Pomeroy *et al.*, 2021; Wilson, 2021), systemic constellations (Ritter & Zamierowski, 2021), and social field pattern development, including social presencing theater work (Gonçalves & Hayashi, 2021), still maintain a difference between participants and researchers, researchers who planned or analyzed alone or chose methodologies for participants. Even Global Social Witnessing (GSW), a contemplative social cognition practice that facilitates mindful witnessing of critical events, is not necessarily participatory and can even be done alone without a community (Matoba, 2021).

3. Ethics of Radical Participatory Design

In order to explicate the ethics, evaluation, and benefits of RPD, I will highlight two projects specifically, while mentioning others. The first was a digital literacy project done under the auspices of a multinational technology company. It was a special project for a vice-president (VP) who wanted a global certification with multiple tracks —a system admin/devOps track, a mobile and web application development track, and a digital literacy track. I will focus on the digital literacy track. Three times, the project failed to reach not just literacy targets but even registration targets. The VP left the organization, and the project lead moved to another project. I was allowed to run the project in any way I chose with the budget. I recruited a team of 12 people mostly from north, central India and participated in an RPD experience to redesign the educational service in a way that would improve digital literacy levels in north-central India to start before expanding to other regions (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). The digital literacy project is an example of a successful RPD project, in which the team experienced sustained and sustainable shifts in power.

I also participated in the redesign of an international summer service-learning program for high school students, in which Washington, DC high school students traveled abroad during the summer doing service-learning projects and then returned home to complete social entrepreneurship projects in DC (Udoewa, 2018, 2022a). This project was completed under the auspices of a nonprofit in collaboration with the local DC public schools district (DCPS). The project to redesign the international summer service-learning program and curriculum is an example of a failed RPD project due to the program's refusal to give up power. I will use both projects to talk about the ethics and evaluation of RPD.

The ethics of general design work apply to RPD, including: confidentiality, anonymity, data disclosures (what, why, and how long data is collected, and when it will be destroyed); transparency and communication of the work and goals; IRB reviews; and research participant referrals when issues come up beyond the skillset and purpose of the designers including trauma issues, etc. Informing the community of the progress, status, and outcomes is also important, though the focus in RPD is informing the wider community since community members are full members of the design team.

Similar to indigenous methodologies, the community leads and decides not just what research is done but also if, what, and how research is shared (Smith, 2021). Usually, RPD communities do not have a preference for the written word. However, when projects are shared in writing, RPD practitioners recognize collaborative and community authorship in two ways. First, all community members who want to co-author a paper can do so (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). Secondly, in non-project papers written alone like this one, I try to cite cultural and community knowledge as a

reference equal to other 3rd-person-knowing, academic author references, not just as a footnote.

Radical Participatory Design requires the addition of remuneration as an ethical concern. It is unjust for a designer to be paid for design work while community members, who are equal designers doing the same work, are unpaid (equality). The injustice is more apparent when we consider that the designer does the work as part of the job while community members must do the work in addition to their normal livelihoods and routines. In cases where community members are jobless or the RPD work takes community members away from their jobs, the offense is greater. Because the design work is not the job of the community member, it costs the community member more to participate in RPD and therefore they should be paid even more (equity).

In the international summer service-learning project, the student community members of the design and research team were not paid for their time. It is possible to say we, the design team, did have equality because the two professional designers, including myself, were also not paid. However, from the standpoint of RPD, ethically, it was still poor practice to fail to compensate the students for their time. Moreover, we did not achieve equity, because failure to compensate design team members had a bigger impact on the students than the impact on the professional designers. The project failed ethically from the standpoint of RPD.

In contrast is the example of a current systems practice RPD project, focused on generational, racialized trauma in the rural U.S. South, the sponsoring nonprofit pays team members (professional designer or community designer) equally according to hours of work. In cases where it is difficult to get approval to pay community members equitably, there are numerous, creative ways to compensate community members. In the digital literacy project in India, I paid for breakfast and lunch each day, a few dinners when it was late in the day, all equipment needed, all travel expenses to work locations, and full room and board for overnight travel and experiential homestay research. I gave references, referrals, recommendations, and certificates of completion to team members to use in job hunting, made the project an internship for resumes, and encouraged team members to publish our work so that they became published authors of two papers.

Community review boards (CRBs) are not a replacement for participation or leadership by the community. They can provide an extra check to prevent unethical, unsafe, inequitable, exclusive research or design from being implemented. Still it is possible for a CRB to become a gatekeeper, setting up a hierarchy filled with the same logics of coloniality. For CRBs to work well, they must be radically participatory and radically representative, like an RPD team. However, they are not a requirement, as an RPD project brings the ethical community checks into the actual research and design process due to the presence of community members. Most RPD projects do not use them.

4. Evaluation of Radical Participatory Design Processes

The examples of remuneration hint at the way to evaluate the success of the RPD process, which is distinct from the success of the design outcome (Drain *et al.*, 2021). An RPD process is successful when a majority of the community designers on the design team experience a sustained and sustainable shift in power. The purpose and goal of RPD is transformation and power exchange. If the power exchange does not occur or is not sustained beyond the work, the RPD process was not successful.

In the international summer service-learning program, I formed an RPD team with students in the program (2018). Though the project had all three RPD characteristics, the organization who initiated the project switched the project to a CPD project when they rejected the student designers' decisions and would not implement them. The students left the experience discouraged, with the same amount of power they had at the beginning of the project. Nothing changed for them. The RPD process was not successful, not radically participatory.

My digital literacy project was the opposite (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). My organization fully relinquished control and implemented what the community designers created. The community owned the narratives of the work and pointed proudly to the outcomes in the news claiming: "We did this. Look what we did!" As a result of the work, they gained experience that helped five of seven community members gain a job. A sixth community member, who was employed, received a promotion. The seventh community member improved his floral business. All became first time authors with two publications. Additionally, our team included three non-designers from within our organization who took a break from their marketing and sales work to do community projects on the ground. All three employees quit their jobs within a year of the experience to focus on similar social impact work because they could not go back, divesting of their power in a multinational company. I did the same. One hundred percent of team members' positions of power at the beginning of the project were transformed and remain that way to this day.

Some of these examples of transformational power exchanges are still within the system of values of those with power, leaving the system unchanged. For example, publishing in a peer-reviewed journal is an achievement that conveys authority and increased power in our current social hierarchy. However, research justice tells us that experiential and cultural knowledge is just as important as published, institutional knowledge. The community may or may not value increased power within an unchanged, oppressive system. Though RPD designers value and fight for increased representation of underutilized communities in traditional seats of power like journal authorship, a higher goal of RPD beyond individual and group power exchanges within the same system is the creation of alternative systems based on community values; this is the goal of pluriversal design. Radical Participatory Design most successfully creates alternative modes of living in the world. Power exchanges for a majority of community members on an RPD design team are still a success because such power exchanges are necessary systemic steps towards pluriversal goals of alternative systems.

5. Benefits of Radical Participatory Design

The benefits of RPD include more successful and effective design outcomes, mutual learning, and power exchanges. Additionally, community members conduct research among other community members. When doing interviews or observation, the familiar community faces help to reduce anxiety. Often, interviewees are more willing to talk and be open with other members of the community. In cross-cultural design and international design projects when professional designers speak a different language from the community, translation is usually needed. In RPD projects, interviews can be conducted in the primary language of the interviewees because design team members speak the language. The ultimate benefit of RPD is embedded local, experiential knowledge in the design team.

6. Beyond Inclusive and Human-centered Design

Radical participatory design facilitates inclusive design and moves beyond it. Instead of only including marginalized community members in research recruitment, RPD places community members as full, equal members of the research and design team. Designers and community members benefit from mutual learning, and the community benefits from a design outcome that is based on their lived, experiential, relational, cultural knowledge. Moving beyond inclusive design, RPD focuses on an inclusive design team. An inclusive team, then, aids in inclusive research recruitment because the team can use their community connections, networks, and lifelong relationships to expedite and facilitate the recruitment process, reducing anxiety more quickly with research participants who recognize the researchers and designers as people from their own community. In the digital literacy project, the design team did not need translators and could go into communities and immediate-

ly reduce anxiety by using the local language and building on networks, connections, and relationships the design team members already had (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). In RPD, the community benefits from better design outcomes due to the inclusion and greater willingness of research participants to offer experiential knowledge and expertise.

The only way to truly achieve a Human-centered Design (HCD) process is through RPD. Human-centered design is a methodology that centers every part of the design process on the community for whom designers are designing. When doing design injustice or CPD, there may be an activity or a phase centered on the community, but the process always moves back to a phase or activity or interactivity work that is not centered on and apart from the community. In an RPD process, every activity of every phase, including interactivity work, is centered on the community because the community is always there, co-leading the design work, driving it forward, and often initiating the work. In cases when the community does not initiate the work, it can still be an RPD approach if the professional designers, who initiate, give up control and power and the community both participates and assumes leadership including leadership that has the power to stop the project.

Ultimately, when conducting RPD, it is common to move away from HCD towards society-centered, community-centered, life-centered, or planet-centered design, methodologies centered on society, community, all life, and the planet at every phase in the process. If a community is truly centered in a design process so radically that they are full-fledged, equal, and equitable team members, then their expertise and desires lead the process. When their expertise leads the process their expertise brings out two dynamics. First, due to the relational nature of existence, to truly care for a group of humans, one must care for the entire ecosystem that nourishes those humans, an ecosystem in which those humans sit. Second, communities care about more than human individuals. They care about their community, society, land stewardship, water resource stewardship systems, etc. Centering the community means centering the cares and priorities of the community which naturally broadens design.

Such a shift benefits not just a specific group, as in HCD, but rather an entire community, society, non-human life like animals and plants, and ecosystems including non-living things such as rivers. Ultimately communities benefit because they have healthier environments and ecosystems, and the design team benefits from learning how to design eco-systematically, relationally, and holistically. Communities near the bottom of social hierarchies tend to be more in tune with the system in which they sit and the various competing needs of both life and non-life in the environment (Gurung, 2020). They bring that knowledge into the process. For example, in a translocal, community design project on water, the community chose this challenge: *Ensure a safe, sustainable, equitable, and affordable drinking water future* (Roberts, 2017). This design challenge is not anthropocentric, but life-centered. In a current, local, community project where we, the design team, are designing a racially just school community, in addition to human needs, we are looking at the building needs, plant needs, compost needs, and more.

One principle of feminist standpoint theory is that people at the bottom of a social hierarchy tend to have a more accurate or holistic picture of reality (Gurung, 2020), which oppressed communities have always known. Since awareness-based methods entered design through Liberatory Design, Equity-centered Design, and more, reflective activities have been added to design processes (Anaissie *et al.*, 2021; Creative Reaction Lab, 2018). However, the focus of awareness-based methods is on people higher in the social hierarchy who do not see parts of the system due to their location in the social field and system. The awareness of the field is generally more communally known to groups at the bottom of the social hierarchy. Instead of expending so much energy to encourage powerful people to reflect on their positionality and the field dynamics, becoming reflexive but not necessarily reflexible —moving towards flexible change— it can be more efficient to simply engage in RPD, shifting the lead-

ership and participation spectra to the community and letting awareness result as a product of the process (Arnold & Schön, 2021).

In contrast to general needs-based design injustice methodologies, RPD brings the benefit of pluriversal design and futures design. Futures design can be considered an asset based approach where the aspirational asset is the shared vision of the future. Pluriversal design seeks to create alternative and multiple modes and ways of being and living in the world according to the values and identities of various communities (Escobar, 2018; Leitão, 2020). Pluriversal design is "a desire-based approach" that opens up the pluriverse, a multiplicity of possibilities, or a world of many worlds which can all be good and different (Escobar, 2018; Leitão, 2020). It is much harder to move from a damage-centered or conventional needs-based approach to a pluriversal, desire-based approach or a future vision, asset-based approach when the community's desires and vision of the future are not represented and voiced in every activity, phase, and interactivity moment of planning and decision-making. Radical Participatory Design provides a platform to converge the desires and visions of the designers and the desires and visions of the community because of the power-exchanging models of "community member as designer" and "designer as community member." This exchange during RPD allows for the visions, desires, values, expertise, and identities of the community to be present and voiced on the design team during the design process, increasing the likelihood that the design process moves to a pluriversal approach, outcome, and a shared vision of the future, if the community is truly leading the process. Thus, RPD is not neutral, but represents a pluriversal bias towards the identities, values, desires, and shared future visions of the community leading the process. For example, in the international summer service-learning project, students mapped out pathways through a future program, not based on problems they experienced, but based on who they wanted to be and what they wanted to become (Udoewa 2018). Thus, in RPD, communities benefit, then, from a design that embodies their local, specific, future vision. This benefit highlights the relationship between decolonization, anticolonialism, and postcolonialism. Decolonization is not the goal. In a postcolonial and neocolonial world, decolonization is the first step on the anticolonial road to a pluralistic multiverse —the pluriverse.

7. Sustained Embodied, Embedded, and Auto-Empathy

One way to hold empathy for community members throughout the entire design process is through RPD. Empathy is one of the primary mindsets and an ideology of HCD (Heylighen & Dong, 2019; Kolawole, 2016). Designers try to gain and keep empathy by researching with community members and carrying the results of that research and the community perspectives into the design stage through qualitative data and design artifacts like personas and empathy maps. But what is empathy? If empathy is understanding and sharing the feelings of another person, we may realize that achieving empathy through a generic design injustice or CPD project is an impossibility.

When viewing empathy through the Global Social Witnessing (GSW) perspective, there are three stages: the witnessing stage in which the observer still feels separate, the sensing stage in which the observer experiences empathy and connectedness with the observed, and the witnessing stage in which the observer experiences oneness with the observed "through mental, affective, and bodily responses" (Matoba, 2021). After the GSW practice, the observer hopefully takes action based on the global empathy gained (Bachen, *et al.*, 2012).

In contrast, Goleman and Ekman identify three components of empathy (Vlismas, 2020). Cognitive empathy is understanding what someone is experiencing, but there still is a distance between the empathizer and the subject of the empathy. Cognitive empathy maps to the observing stage of GSW and the research stage of HCD. The second component of empathy is emotional empathy. Emotional empathy is feeling with someone, experiencing the same feelings and sharing in that experience. The empathizer has now put themself

in the same emotional space as the subject of the empathy, walking alongside the subject through their emotional journey. Emotional empathy can extend to physical sensations as well, and maps to the sensing stage of GSW and the synthesis and define stage of HCD. Lastly, there is compassionate empathy. Compassionate empathy is being moved to help. It is a balance between cognitive and emotional empathy, where the empathizer is not overwhelmed and paralyzed by emotion (emotional empathy) and simultaneously does not immediately jump into problem solving based on understanding (cognitive empathy). Compassionate empathy maps to the last witnessing stage of GSW, in which the observer experiences oneness with the observed through responses, and to the design and delivery stage of HCD.

When we understand empathy, not as one type or another, but as the summation, co-mingling, or relation of all empathic components, we know that empathy is not required for designers to engage in a design injustice or CPD project. Most designers work with an intellectual understanding of community members' experiences, and then work to change the situation or design a solution. It is clear that one component, cognitive empathy, can be temporarily achieved through research. The difficulty is maintaining the cognitive empathy initially achieved, and achieving emotional empathy and compassionate empathy.

In work with experienced senior designers, cognitive bias slips into the design process blocking cognitive empathy. The further away in time designers are from the research that informs the design, the less cognitive empathy the design team has. Cognitive biases even appear directly after research in the awareness-based sense-making and synthesis phase, when experienced designers and design researchers make claims or extract insights that are not based on patterns but rather the last piece of information they read, the most recent interview debrief, or one interview, observation, activity, or report that they remembered quite well. When I ask what data the claim is based on, I discover that it is scant or not there. Even artifacts, like personas, that are meant to carry cognitive empathy into later stages of the design process can falter due to various reasons: irrelevant information included in personas that designers implicitly and cognitively interpret as important, persona photos or images whose demographics and physical appearances are erroneously associated with subcommunity members introducing more bias, obsolete personas which are incorrectly treated as current because designers do not continuously update them, and the complete lack of use of a persona in the design process after its creation as if the simple act of creation is enough to generate empathetic fitness or empathetic endurance (Farai, 2020). The same analysis can be applied to other design artifacts like empathy maps, days-in-the-life, etc.

In addition to the fleeting nature of cognitive empathy, it is rare for designers to create emotional empathy. Because it is not required in the design process, it is not measured, captured, or evaluated. There may be designers who achieve it and others who do not. Anecdotally, emotional empathy is rare from my experience; most designers are referring to cognitive empathy when they use the term empathy. Additionally, a large barrier to emotional empathy is the lack of sufficient relational time in the context of the power hierarchy between the designer and research participant. Though designers could utilize more longitudinal studies interacting with the same participant over time, most design studies involve a single interaction with a community member during a research phase. Compare a single design interview to the repeated interaction over months that a clinical psychologist or therapist has with a patient. Even in the therapeutic context, MacNaughton (2009) argues that empathy is impossible due to the imbalance in the relationship. Over time, the building of relationship and psychological trust can reduce or temporarily suspend the power imbalance enough to allow the possibility of empathic transfer; however a single design research interview is insufficient to achieve this. Another obstacle to emotional empathy is the lack of experiential research in many projects. It is difficult to gain emotional empathy through interviews alone without actual experiences. Simulations and experiential methods like mystery shopping, mys-

tery working, homestays, participant observation, work-alongs, etc. are much more powerful at evoking or provoking designers emotionally to move towards emotional empathy (Stickdorn *et al.*, 2018; Woodcock *et al.*, 2019). However, they are not used as much as the interview method. Finally, what happens when one has a strong aversion to the lifestyle or values of the community members one is researching? One might have cognitive empathy but emotional empathy may be an impossibility due to conflicting values and worldviews. This situation leads to critical empathy which is not emotional empathy (de Coning, 2021).

However, designers do respond to the community needs which would seemingly qualify as compassionate empathy. The difficulty is that compassionate empathy is not simply responding to help or acting, it is being moved to respond to help, and then helping. Compassionate empathy necessarily requires emotional empathy. Even if emotional empathy were not required, the motivation for compassionate empathy must be compassion. It is impossible to achieve this in the case of a professional designer because the financial incentives, wages, or salary make such compassionate motivation impossible. With or without compassion, the designer's job and goal is to act and receive compensation. Even when designers conduct pro bono or volunteer work, the framing of the work or the agreement is that the designer will conduct research, uncover important insights, and create something. This framing or prior agreement makes compassionate empathy an impossibility. We can never know if the designer would have been moved to act and then act, outside of an agreement that dictates they will act.

If not an impossibility, empathy is rare (Nathanson, 2003; Macnaughton, 2009; Watson, 2009). How can we ever truly, experientially know what someone else is going through (Heylighen & Dong, 2019)? It is much better to avoid the problem of gaining empathy. For example, in the international summer service-learning project, instead of the designers building personas to create empathy, the students built auto-personas of themselves (Udoewa, 2018). Radical participatory design avoids the problem of gaining empathy by simply embedding empathy through lived, communal, embodied, cultural, and spiritual experience and experienced emotional journeys, into the design team for the entirety of the design process. Instead of relying on transcripts and research artifacts to create empathy and hold the community needs in the forefront of the minds of the designers throughout implementation, the presence of community members on the team brings their lived experience into all conversations, decisions, explanations, and implementations. That lived experience can check a process, encourage, cajole, explain, remind, expand, teach, and familiarize. This converts the design process not only into a power exchange but also an emotional exchange between team members as the professional designers on the team relate to, engage with, connect to, and learn from their community member teammates and designers. Such an empathic exchange benefits communities by creating a design outcome fully driven by and embedded in their experience. Designers benefit through mutual learning and the gift of relationship.

8. Systemic Action

Radical participatory design has a higher likelihood of creating systemic action or active non-action (refusing to act unjustly) than other awareness-based methodologies. While awareness-based methodologies and methods can and have led to some kind of action, two difficulties with action arise. First, in my lived experience, often they do not lead to action. This is due to the fact that post-awareness action is always a choice. Similar to the bystander in GSW or the "white moderates" to whom Dr. King wrote from prison, a person can become aware of a situation or the plight of another and choose not to act due to fear, very high costs, fragility, system-reinforcing punishment of anti-systemic behavior, etc. (DiAngelo, 2018; King, 2018; Matoba, 2021). There can be a disconnection and stagnant absencing between the presencing steps of open heart and open will (Scharmer, 2009). In some situations, people who are aware of social injustice pretend to be unaware or do not acknowledge it at the conscious level even when their subconscious knows

it to be true (Gilson, 2022; Pohlhaus, 2012). Pomeroy *et al.* ask: "what are the methods that best serve action" as an open question (2021, p. 115). Radical participatory design moves from awareness to social action; it is a design meta-methodology. Action is often a result of design processes that implement something. Radical participatory design goes further because it moves to social action. Cunningham reflects that awareness-based methodologies "don't unmake centuries of injustice and violence by being generative in a room, but [they] do help the social body in the room become more effective at the thing they are trying to do" (Cunningham, 2021, p. 12). Awareness-based methods are more focused on making people more effective at the work they are already doing with some level of awareness, while RPD actually creates new work and actions by the very nature of being a design meta-methodology. The new work can be considered a trivial outcome because design, by definition, usually creates new things. Still, any system-oriented design, especially one like RPD that changes the structure of relationships and connections in a system, has an advantage over awareness processes that may not lead to new work or actions. RPD creates new actions and work for professional designers who may be completely unaware, thrust into an environment of relational knowledge, or for newly self-empowered community members due to the active divestment of power by professional designers. Second, often contemplative and awareness-based methods lead to personal or insular change and never transition to change for social justice. I had this conversation with participants while participating in a contemplative dance workshop that moves from emotions to art, from art to awareness, and from awareness to action. Often the change or action is personal and there can be a disconnection to larger, needed social change and actions. Because of the shift of the leadership spectrum to the community, RPD often leads to social change, evidenced in various movements such as U.S. civil rights and labor rights (Udoewa, 2022a, 2022b in press).

Radical participatory design tends to create more systemic action by inviting new entrants into systems change. As Cunningham notes, awareness-based methods improve the effectiveness of what change agents are already trying to do (2021). They do not necessarily invite more people into the work. If the fundamental work required for systems change is to align the purposes and awareness of all system actors, awareness needs to spread to people who do not practice awareness-based methods. However, there are people who are not willing to participate in contemplative or awareness-based practices as they are not comfortable or accustomed to operating from that emotional or spiritual center. However, participatory design is often defined and viewed as a participatory way of practicing design, and a person may not realize the awareness-based dimensions of the practice. Designers new to systems change may practice RPD with less trepidation than an explicit awareness-based practice.

Lastly, RPD more naturally leads to systems practice (a practice focused on improving the health of a system), futures design (the use of longer-term forecasting or visioning to drive design choices in the present), and other asset-based methodologies. Due to the shift of the leadership spectrum to the community, the community chooses the methodology instead of the professional designer, opening up a variety of possibilities. When conducting RPD, it is quite natural for the work to become asset-based because community members naturally define themselves by what they offer and what gifts they bring, not by what they lack. I view systems practice and futures design as asset-based methodologies because instead of focusing on the problem, they focus on assets: the system dynamics and health, and a shared vision of the future, respectively. Community members know, implicitly or explicitly, the dynamics of the system in which they sit and often highlight the system concerns and the interconnectedness of the system components when the design team is considering the plausibility of a particular option. Communities contain deep experiential and cultural wisdom that understands the system and underscores needs outside of human needs. For example, I work on a community project where the team has designed several system interventions based on high-impact leverage points found while

analyzing a system map the team created based on their systems research of generational racialized trauma in the rural south (Jagannathan & Seugling, 2018). Through RPD, communities, society, and the environment benefit from more systemic solutions, avoiding HCD solutions that leave the problem unaddressed, make it worse, or only temporarily resolve it. Designers benefit from learning systems practice skills.

9. Trauma-responsive Design

Radical participatory design is a more effective approach to practice trauma-informed and trauma-responsive design than trauma-informed design based only on mainstream institutional knowledge (Jackson *et al.*, 2020). All designers, including RPD designers, should practice trauma-informed design because it is not possible to know if a particular community member, interacting with researchers or their designs, has experienced trauma. One 2016 epidemiological study, conducted in twenty-four countries, found that over 70% of research respondents had experienced at least one trauma event and 30.5 per cent had experienced four or more trauma events (Benjet *et al.*, 2015). When working among historically and presently marginalized, colonized, and oppressed communities, the percentage of people experiencing trauma can be even higher.

Trauma-informed design is design that involves three components. First, trauma-informed design is design that recognizes that people can have many different traumas in their lives including past traumatic events whose adverse effects can still be present today, as well as the possible paths to recovery. Second, trauma-informed design involves designers who recognize the signs and symptoms of trauma in participants, researchers, and societal systems. Third, and most importantly, trauma-informed design is designing in ways to avoid triggering and to resist retraumatizing participants through research and design work and interactions.

The US Substance Abuse and Mental Health Services Administration (SAMHSA) defines six principles of a trauma-informed approach (SAMHSA, 2014).

1. Safety.
2. Trustworthiness and Transparency.
3. Peer Support.
4. Collaboration and Mutuality.
5. Empowerment, Voice and Choice.
6. Cultural, Historical, and Gender Issues.

Instead of the designer having to carry the weight of ensuring these six principles, RPD bypasses this work. In the RPD approach, peer support, collaboration, mutuality, empowerment, voice, and choice are more naturally a part of the process because the community is participating, leading, and driving the process. The community brings its cultural and historical knowledge and lived experience including gender identities and issues. Because their presence is welcome and their voices are heard and they see other community members leading, safety is increased and anxiety is reduced both in the research and design process and in interacting with designs created by the design team. Trust is increased and community design members offer transparency and communication to the broader community about the work they are doing. In other words, the SAMHSA principles are built into the RPD framework naturally as the community is embedded on the design team as equal, full members with leading voices.

Of course, in general life, traumatized people can traumatize other people. In a design process, a trauma-informed design team may be at a disadvantage if their practices are only based on mainstream, institutional, social work knowledge. Because mainstream, institutional knowledge or 3rd-person knowing is studying lives, bodies, experiences, cultures, communities, and more, it is always behind the lived, experiential, embodied, intuitive, relational, communal, cultural, and spiritual knowledge itself. Through the RPD meta-methodology, the team is better able to be trauma-responsive due to a greater array of epistemologies providing and embedding more current trauma information and updated care practices. For instance, in the digital literacy project, the design team was able to avoid triggering and retraumatizing often forgotten people in temporary housing because the design team was composed of community members who knew

what it was like to be in such a situation (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). Through RPD, designers benefit from experiential knowledge-based and skill-based trauma-responsive practices, and community members and designers benefit from reduced triggers and harm, as well as an increased sense of care and belonging.

10. Challenges and Barriers to Radical Participatory Design

The fundamental and most dangerous difficulty of RPD is the tendency for an RPD process to stop being critically and radically participatory. This can be mediated by the second characteristic of RDP: community members outnumber the designers.

In designer-initiated projects, an RPD process may flip to CPD, as well, because the designer or the design organization decides to usurp or regain control, rejecting the work of the community. Perhaps the community members were only invited for a short period of time so the project reverts back to an organizational design injustice process. Or the designers and the organization never made plans for the critical involvement past the design phase into the implementation phase. To address these pitfalls, I have learned several lessons from experience. Strategically work to institutionalize RPD in the organization so that an RPD project is not simply a one-time event or an approval process each time. Secure resources, such as funding, to make RPD a continual part of the project work. Contract community members throughout the lifecycle of the product or service. Create transition plans so that community members can retire from the RPD work and new community members can join and take the place of the retiring members. Practice relinquishing power daily. Due to the structure of society and the continued aggregation and consolidation of power in certain organizations and people, it is important that the divestment of power be a continual practice. The designers and the design organization should continue to divest of power while the community members assume power even in the implementation stage. When the divestment of power is done to the core, even if a design organization wanted to take control and run the project differently, the organization could not do this. The design artifacts, the narratives, and the resources are all within the control of community members. If these resources are not within the control of the community, the designers and design organization did not truly give up all power. In the following section, I will give advice on how to choose projects where the organization is more likely to give up power.

Another challenge for designers in the RPD process is privileging the process and their expertise over the lived experience of the community members. The nonlinearity of the process should not come from a designer's power as facilitator, making decisions and planning apart from the community between design activities. The nonlinearity in the design process should come from the insight, inclination, needs, desires, and even disruption of the community member designers on the team (Knutz & Markussen, 2020). In the recent digital literacy project, the design team implemented a positive deviance research method in the middle of prototyping because the community member designers were feeling uninspired (Udoewa *et al.*, 2016; Udoewa *et al.*, 2017). Likewise, in the same project, I, a professional designer, did not "correct" their designs from a Western Anglocentric design perspective, but watched them aesthetically design what was most pleasing to them, based on their experiences and values. We were practicing pluriversal design.

Keeping in mind that the design team may be more heterogeneous as a result of the RPD process, one must pay more attention to team dynamics (Huybrechts *et al.*, 2020). Mixing community members with designers of largely homogenous backgrounds juxtaposes multiple subcultures together. The team must work to establish a strong foundation of trust and safety, and then, upon that foundation, cultivate a culture that mines for conflicting ideas in order to get to the best ideas (Lencioni, 2012). This type of culture is not automatic and must be built on any team, especially and including an RPD team

which may have designers who have never worked with community members and vice versa.

Making decisions in ways that do not privilege the designers can be difficult. There is no one way RPD teams make decisions because, generally, design team members try to use culturally appropriate ways of making decisions. Usually, we, design team members, decide as a group how to make decisions in such a way that everyone will support the decision, even if the decision was not their personal choice. In order for the support to be present when decisions are made, we decide how to decide, using either unanimity, consensus, or consent-based decision making (Bockelbrink *et al.*, 2022). Once a particular choice is unanimous or we have a consensus or complete consent, we can proceed to make decisions using the chosen decision making process. In some RPD, there is also an eco-relational approach to the politics of decision making in which people do not voice individual desires but simply carry out tasks with aligned purpose, like parts of the human body. The ecological system of people makes decisions based on the collective purpose (similar to the way that blood might rush the limbs during a flight, fight, or freeze response without any part of the body making an explicit conscious decision).

Due to the educational nature of the RPD process, decolonial concepts of time, and lives of community members, RPD may take longer than design injustice or CPD because of availability, pace of community life, decentering white-supremacist sense of urgency, and the many learning and practicing sessions (Smith, 2021; Mowris, 2020; Creative Reaction Labs, n.d.). When compared to CPD outcomes over shorter project timelines, communities alongside whom I have worked value the RPD outcomes over the longer time. It is helpful to plan for this time and flexibility from the start and communicate the flexibility and timelines to stakeholders and community members.

Lastly, RPD does not avoid the problem of bias on the design team. In fact, the participating community members may represent a biased portion of the community and their biased lived experience can shift the work the design team does, creating designs that do not serve other portions of the community (Taoka *et al.*, 2018). To counteract this effect, choose a qualitatively representative sample, when possible. Avoid looking for a representative from every family in a community or subgroup. Rather, list all the attributes of community members that might alter how one would design for them. Then make sure the design team has community members from different parts of each attribute spectrum (IDEO.org, 2016). Any bias or limited knowledge on the design team should be addressed by recruiting a qualitatively representative sample of the community as research participants. The bias of the community members on the design team can still affect the process. Conduct "Beginner's Mindset," "Observing vs Interpreting," and other bias awareness training like bias journaling for the entire team (IDEO.org, 2007). I usually repeat bias journaling weekly and review my writings ahead of each research session. Ideally, conduct the training sessions before the research and interview and observation guide creation. The training does not eliminate bias, but serves to make the entire design team more aware of their bias and, thereby, to limit its adverse impact.

Organizations can still pose a barrier. It is unnatural for those with power to surrender it, a requirement for the success of the RPD process which involves a power exchange. Organizational leaders often prohibit RPD work because they do not want to invest the time or resources. Others do not want to invest in proper ethical treatment of external community members. Others do not understand what purpose designers have if design can be done by anyone. Many are afraid of anything new, and are change-averse. If the organization and its methods are successful by some measure, they do not want to change it. Others do not trust community members and want to retain control.

11. Conclusion

There are many challenges when participating in RPD work. The design team must take care to plan for a longer, educational process, working to reduce

bias on the design team, and specifically working to prevent the RPD process from switching to a CPD process. The act of divesting of power is a continual act into which the designers and design organization must repeatedly enter. Ultimately, an RPD process is most successful when alternative systems of value and ways of living in the world are created.

Organizations resist giving up power. One barrier is not understanding the purpose of designers and paying for design services if the community can design. Designers have honed a craft that can be helpful to the RPD process. Their knowledge is not privileged above community experiential and cultural knowledge. Designer knowledge is still useful and especially powerful when combined with community knowledge. For example, a community, practicing community-driven design, might call a structural engineer to validate their building design. I have worked on a learning design project in which the community unearthed learning design principles and created designs based on learner needs without a learning designer. Still, if needed, communities may invite designers at any stage such as research, design, or implementation.

Ultimately, the invitation to divest of power, as a designer, can still be accepted even if the organization refuses to do so. One can divest of one's power by leaving such organizations. The best way to engage in RPD work is not to fight unwilling organizations, though important, but to work with fellow community members in the local community on local problems. This work will automatically be RPD because one is a community member, not an outside designer. The designer's design skills are a benefit to the community just as the skills of the other members are a benefit to the community. Foremost, the designer's lived experience in the community makes the designer a member and positioned to co-lead and drive the work alongside other members.

In future work, I will go beyond general relational design which includes RPD and elaborate on a subset of RPD that I call Relational Design. In Relational Design, design team members do not only design relationally, or alongside community members. Design team members also replace various extractive and transactional steps in the generic design process with explicitly dialogic and relationship methodologies and activities. Secondly, I want to elaborate more on the decision-making process and options in an RPD project. I will show what RPD decision-making looks like, highlight a relational and biosystems approach to decision making, and share how to make decisions in a way to minimize the likelihood that an RPD project flips to a CPD project.

12. References

Anaissie, T., Cary, V., Clifford, D., Malarkey, T., & Wise, S. (2021). Liberatory design.

Arnold, R., & Schön, M. (2021). The reflexible person: Toward an epistemological learning culture. *Journal of Awareness-Based Systems Change, 1*(2), 51-71. https://doi.org/10.47061/jabsc.v1i2.971

Bachen, C., Hernández-Ramos, P. F., & Raphael, C. (2012). Simulating real lives: Promoting global empathy and interest in learning through simulation games. *Simulation & Gaming, 43*(4), 437-460. https://doi.org/10.1177/1046878111432108

Benjet, C., Bromet, E., Karam, E. G., Kessler, R. C., McLaughlin, K. A., Ruscio, A. M., Shahly, V., Stein, D. J., Petukhova, M., Hill, E., Alonso, J., Atwoli, L., Bunting, B., Bruffaerts, R., Caldas-de-Almeida, J. M., de Girolamo, G., Florescu, S., Gureje, O., Huang, Y., Lepine, … Koenen, K. C. (2015). The epidemiology of traumatic event exposure worldwide: Results from the World Mental Health Survey Consortium. *Psychological Medicine, 46*(2), 327-343. https://doi.org/10.1017/s0033291715001981

Bockelbrink, B., Priest, J., & David, L. (2022, April 24). *Consent Decision-Making*. A Practical Guide to Sociocracy 3.0. https://patterns.sociocracy30.org/consent-decision-making.html

de Coning, A. (2021). Seven theses on critical empathy: A methodological framework for "unsavory" populations. *Qualitative Research*. https://doi.org/10.1177/14687941211019563

Cooke, B., & Kothari, U. (Eds.). (2001). *Participation: The new tyranny?* Zed Books.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press.

Creative Reaction Lab. (2018). *Equity-centered community design field guide.* Retrieved from https://static1.squarespace.com/static/5e3b20447d777f2b32c1bc1c/t/5e667103feb2830f1b1b68d4/1583771908636/ECCD+FIELD+GUIDE+-+DOWNLOAD.pdf

Creative Reaction Lab (n.d.). *How Design Thinking Protects White Supremacy.* Creative Reaction Lab. https://crxlab.org/webinar-series

Cunningham, D. (2021). Awareness based systems change and racial justice. *Journal of Awareness-Based Systems Change, 1*(2), 9-13. https://doi.org/10.47061/jabsc.v1i2.1590

DiAngelo, R. (2018). *White fragility: Why it's so hard for white people to talk about racism*. Beacon Press.

Drain, A., Shekar, A., & Grigg, N. (2021). Insights, solutions and empowerment: A framework for evaluating participatory design. *CoDesign, 17*(1), 1-21. https://doi.org/10.1080/15710882.2018.1540641

Ellison, C. (2014). *Indigenous knowledge and knowledge synthesis, translation and exchange (KSTE)*. National Collaborating Center for Aboriginal Health.

Escobar, A. (2018). Designs for the Pluriverse. Duke University Press. https://doi.org/10.1215/ 9780822371816

Farai, S. (2022, November). The Impossibility and Irrelevance of Empathy. Lecture. https://joinlearners.com/talk/the-impossibility-and-irrelevance-of-empathy

Geppert, A. A., & Forlano, L. E. (2022). Design for equivalence: Agonism for collective emancipation in participatory design. *Participatory Design Conference 2022, 1*, pp. 158-168. https://doi.org/10.1145/3536169.3537790

Gilson, E. (2022) Sexual injustice and willful ignorance. In M. Gross & L. McGoey (Eds.), *Routledge International Handbook of Ignorance Studies* (pp. 257-268). Routledge. https://doi.org/10.4324/9781003100607-29

Gonçalves, R. D., & Hayashi, A. (2021). A pattern language for social field shifts: cultivating embodied and perceptual capacities of social groups through aesthetics, and social field archetypes. *Journal of Awareness-Based Systems Change, 1*(1), 35-57. https://doi.org/10.47061/jabsc.v1i1.478

Gurung, L. (2020). Feminist standpoint theory: Conceptualization and utility. *Dhaulagiri Journal of Sociology and Anthropology, 14*, 106-115. https://doi.org/10.3126/dsaj.v14i0.27357

Heylighen, A., & Dong, A. (2019). To empathise or not to empathise? Empathy and its limits in design. *Design Studies, 65*, 107-124. https://doi.org/10.1016/j.destud.2019.10.007

Hill, D. (2012). *Dark matter and trojan horses: A strategic design vocabulary*. Strelka.

Huybrechts, L., Teli, M., Zuljevic, M., & Bettega, M. (2020). Visions that change. Articulating the politics of participatory design. *CoDesign, 16*(1), 3-16. https://doi.org/10.1080/15710882.2020.1728907

IDEO.org. (2007). *Human-centered design toolkit*. San Francisco. https://www.ideo.com/post/design-kit

IDEO.org. (2016). *Extremes and mainstream. Design Kit*. https://www.designkit.org/methods/45.

Ibibio. (Generations). Nam means make, create, or design in Ibibio.

Jackson, D., Kim, M., & Sievert, J. R. (2020). The rapid embrace of legal design and the use of co-design to avoid enshrining systemic bias. *Design Issues, 36*(3), 16-30. https://doi.org/10.1162/desi_a_00601

Jagannathan, V., & Seugling, S. (2018, March 7). A systems map of generational trauma in the rural South. Kumu. https://bailey.kumu.io/a-systems-map-of-generational-trauma-in-the-rural-south

King, M. L., Jr. (2018). Letter from Birmingham jail. Penguin Books.

Knutz, E., & Markussen, T. (2020). Politics of participation in design research: Learning from participatory art. *Design Issues, 36*(1), 59-76.

Kohtala, C., Hyysalo, S., & Whalen, J. (2020). A taxonomy of users' active design engagement in the 21st century. *Design Studies, 67*, 27-54. https://doi.org/10.1016/j.destud.2019.11.008

Kolawole, E. (2016). Empathy. *Design Kit*. https://www.designkit.org/mindsets/4.

Law, J. (2015). What's wrong with a one-world world? *Distinktion: Scandinavian Journal of Social Theory, 16*(1), 126-139. https://doi.org/10.1080/1600910x.2015.1020066

Leitão, R. (2020). Pluriversal design and desire-based design: Desire as the impulse for human flourishing. In R. Leitão, L. Noel, & L. Murphy (Eds.), *Pivot 2020 Conference Proceedings*. https://doi.org/10.21606/pluriversal.2020.011

Lencioni, P. M. (2012). *The five dysfunctions of a team: Team assessment*. John Wiley & Sons.

Macnaughton, J. (2009). The dangerous practice of empathy. *The Lancet, 373*(9679), 1940-1941. https://doi.org/10.1016/S0140-6736(09)61055-2

Matoba, K. (2021). Global social witnessing: An educational tool for awareness-based systems change in the era of global humanitarian and planetary crisis. *Journal of Awareness-Based Systems Change, 1*(1), 59-74. https://doi.org/10.47061/jabsc.v1i1.548

Mowris, L. G. (2020). DISMANTLING WHITE SUPREMACY CULTURE: developing shared language for the poison and the antidote. Atlanta, GA; lenspeace.

Nathanson, A. I. (2003). Rethinking empathy. In J. Bryant, D. R. Roskos-Ewoldsen & J. Cantor (Eds.), *Communication and emotion* (pp. 115-138). Routledge. https://doi.org/10.4324/9781410607584-11

Pohlhaus, G. (2012). Relational knowing and epistemic injustice: Toward a theory of willful hermeneutical ignorance. *Hypatia, 27*(4), 715-735. https://doi.org/10.1111/j.1527-2001.2011.01222.x

Pomeroy, E., Herrmann, L., Jung, S., Laenens, E., Pastorini, L., & Ruiter, A. (2021). Exploring action research from a social field perspective. *Journal of Awareness-Based Systems Change, 1*(1), 105-117. https://doi.org/10.47061/jabsc.v1i1.676

Peter, R. (2006). *How (not) to speak of God: Marks of the emerging church*. Paraclete Press.

Ritter, L., & Zamierowski, N. (2021). Systems sensing and systemic constellation for organizational transformation: Building collective capacity for navigating complexity. *Journal of Awareness-Based Systems Change, 1*(2), 101-115. https://doi.org/10.47061/jabsc.v1i2.1181

Roberts, M. (2017). *Emerging leaders in science and society: 2016 annual report.* American Association for the Advancement of Science. https://www.aaas.org/sites/default/files/ELISS-Annual-Report-2016-1.pdf

Rollins, P. (2006). How (not) to speak of God (p. 144). Brewster, MA: Paraclete Press.

Substance Abuse and Mental Health Services Administration. (2014). *SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach*. HHS Publication No.(SMA) 14-4884. https://ncsacw.samhsa.gov/userfiles/files/SAMHSA_Trauma.pdf.

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *Co-design, 4*(1), 5-18. https://doi.org/10.1080/15710880701875068

Suaalii-Sauni, T., & Fulu-Aiolupotea, S. M. (2014). Decolonising Pacific research, building Pacific research communities and de-

veloping Pacific research tools: The case of the talanoa and the faafaletui in Samoa. *Asia Pacific Viewpoint, 55*(3), 331-344. https://doi.org/10.1111/apv.12061

Sbardelini, R., Almeida, D., & Ramos, L. M. (2022). MAPA: Co-creating new narratives for the 21st century. *Journal of Awareness-Based Systems Change, 2*(1), 101-118. https://doi.org/10.47061/jabsc.v2i1.2727

Scharmer, C. O. (2009). *Theory U: Learning from the future as it emerges*. Berrett-Koehler Publishers.

Smith, L. T. (2021). *Decolonizing methodologies: Research and indigenous peoples*. Bloomsbury Publishing.

Stickdorn, M., Hormess, M. E., Lawrence, A., & Schneider, J. (2018). *This is service design doing: Applying service design thinking in the real world*. O'Reilly Media.

Taoka, Y., Kagohashi, K., & Mougenot, C. (2018). A cross-cultural study of co-design: The impact of power distance on group dynamics in Japan. *CoDesign*, *17*(1), 22-49. https://doi.org/10.1080/15710882.2018.1546321

Udoewa, V., Mathew, N., Al-Hafidh, L., Bhog, L., Gupta, A., Patel, P., Prabhakar, B., Kaushik, S., Bauer, L., & Humar, V. (2016). Helping the next 4 billion go online part I: Design research for digital literacy education. *International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship, 11*(2), 18-37. https://doi.org/10.24908/ijsle.v11i2.6393

Udoewa, V., Mathew, N., Gupta, A., Bauer, L., Bhog, L., Prabhakar, B., Patel, P. & Al-Hafidh, L. (2017). Helping the next 4 billion go online part II: Prototyping solutions for digital literacy education. *International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship, 12*(1), 13-40. https://doi.org/10.24908/ijsle.v12i1.6665

Udoewa, V. (2018). Redesign of a service-learning social entrepreneurship program for high school students Part I. *International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship, 13*(2), 79-92. https://doi.org/10.24908/ijsle.v13i2.11492

Udoewa, V. (2022a). Hybrid expeditionary service learning model. *International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship, 17*(1), 20-36. https://doi.org/10.24908/ijsle.v17i1.14810

Udoewa, V. (2022b in press). An Introduction to radical participatory design: decolonizing participatory design processes. *Design Science*, *8*.

Vlismas, T. (2020, June 24). What is empathy? Learn about 3 types of empathy [web log]. Retrieved November 23, 2022, from https://takealtus.com/2020/06/empathy-1/#:~:text=Renowned%20psychologists%20Daniel%20Goleman%20and,%3A%20Cognitive%2C%20Emotional%20and%20Compassionate.

Watson, C. (2009). The 'impossible vanity': Uses and abuses of empathy in qualitative inquiry. *Qualitative Research, 9*(1), 105-117. https://doi.org/10.1177/1468794108098033

Wilson, P. (2021). Sensing the Social Field through Action Research: What's important, what's valid. *Journal of Awareness-Based Systems Change, 1*(1), 119-124. https://doi.org/10.47061/jabsc.v1i1.679

Woodcock, A., Osmond, J., Tovey, M., & McDonagh, D. (2019). Empathy thresholds in transport design students. *Design and Technology Education, 24*(1).

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its overrepresentation-An argument. *CR: The new centennial review, 3*(3), 257-337. https://doi.org/10.1353/ncr.2004.0015

Wynter, S., & McKittrick, K. (2015). Unparalleled catastrophe for our species? Or, to give humanness a different future: Conversations. In K. McKittrick (Ed.), *Sylvia Wynter: On being human as praxis* (pp. 9-89). Duke University Press. https://doi.org/10.1515/9780822375852-003

Hernaniko Kulturarteko Plaza Feminista. Iturria: Itziar Bastarrika - Hernaniko Udala

Ikuspegi feminista eta integratzailea ezinbestekoa da Hernaniko Emakumeen Etxearen proiektuan

Sukaldaritzaren esanahi soziala eta kulturala ikuspegi feministatik azter daiteke. Ikerketa-talde batek azterketa bat egin du Hernaniko Emakumeen Etxea proiektuaren testuinguruan, Euskal Herrian. Bertan, sukalde bat sartzeari buruzko eztabaidan sakondu da. Nabarmendu da kideek parte hartzeko prozesu bat abiatu dutela barne-diseinua eta etxearen funtzionamendua zehazteko. Sukaldea fazeta anitzeko gune gisa islatzen da, dimentsio politiko, sinboliko eta afektiboak irudikatuz.

Emakumeen Etxea mugimendu feministen eta arrazismoaren aurkako gune gisa erabiltzen da, eta, beraz, beharrezkoa da gizarte-aniztasuna kudeatzea eta elkartearen barruan adostasuna lortzea. Sukaldea tentsio kritiko eta sortzaileko leku gisa aurkezten da metaforikoki, eta ikerketa, irudimena eta eraikuntza sustatzen dira.

Euskal gizartean gertatzen diren aldaketa sozial zabalenak eta politika feministaren izaera ebolutiboa aztertu behar dira. Sukaldearen eta etxeko eremuaren inguruko alderdi teoriko eta praktikoak izaki, kulturen arteko topaketek eraldatze-ahalmena dute, eta ikuspegi feministek araumugei aurre egin eta zenbait subjektibotasun politiko sustatu ditzakete.

Ikertzaileek hainbat gizarte-mugimendu aztertu dituzte, batez ere feminismoa, arrazismoaren aurkako mugimendua, ekologismoa eta euskararen sustapena. Metodo kualitatibo eta etnografikoak erabiliz, landa-lana eta elkarrizketak barne, Hernaniko Emakumeen Etxea Gipuzkoako beste emakume-etxe batzuekin konparatu dute, eta gizarte-eraldaketen eragina nabarmentzen dute, hala nola desinstituzionalizazioa, ekologia- eta laguntza-krisiak eta desberdintasunen hazkundea. Feminismoak funtsezko zeregina du euskal gizarte-mugimenduan, parte-hartze dinamikoaren eta sektoreen arteko ikuspegien alde eginez. Etxea baliagarri izango da gizartea eraldatzeko eta esperimentatzeko laborategi gisa.

Sukaldearen kontzeptua aztergai

Euskal kulturan, historikoki, emakumeak elkarte gastronomikoetatik baztertu izan dira, eta, hortaz, ikuspegi feministak eztabaidatu dira. Ikertzaileek Emakumeen Etxean sukaldeak duen garrantzia aztertu ondoren, jarrera tradizionalak desikastearen garrantzia azpimarratu dute. Emakumeen Etxeak elkartasun komunitarioko modu berriak probatzeko laborategiak dira. Testuinguru horretan, emakume etorkinen erronkak ere aztertu dira, baita espazio feministen barruan ere.

Azkenean, ikertzaileek parte-hartzearen garrantzia azpimarratu dute. Ezinbestekoa da emakumeek erabakiak hartzeko prozesuetan parte hartzea; bereziki, erakundeen eta gizarte-mugimenduen arteko lankidetzak eskatzen dituzten proiektuetan.

Emakumeen Etxea sortzeko prozesuaren barruan, oztopo handiak sortu ziren sukaldeko eztabaidan. Horregatik, ikertzaileek elkarrizketaren eta ikuspegi zirkularraren garrantzia azpimarratzen dute. Ikertzaileek hausnartu dute Emakumeen Etxearen proiektuaren ikuspegi feministak etengabe ebaluatzeak eta zabaltzeak duen garrantziari buruz. Azken finean, sukaldaritza prozesu kolektibo baten sinboloa da, eta Emakumeen Etxean egindako eztabaidak, lortutako akordioak eta politika feministak dakartzan konplexutasunak eta aldaketak erakusten ditu.

InpAkta 00 (2023) 188-202

https://doi.org/10.1177/14661381211073122

Ethnography of the kitchen: the women's house, a space for feminist alliance and intercultural encounter

Mari-Luz Esteban*, Miren Guilló-Arakistain, Marta Luxán-Serrano

AFIT-Feminist Anthropology Research Group, University of the Basque Country, UPV/EHU

ABSTRACT: In this article we delve into a debate about whether a kitchen was to be installed in a new Women's House in a city in the Basque Country (Hernani, Gipuzkoa). The ethnography presented here was conducted by observing the process around the creation of the House. Articulating the debate's main points led us to examine the dominant cultural assumptions about cooking in Basque society, especially in view of the opposing feminist positions on the kitchen and the domestic sphere. To understand the changes that took place, it is essential to consider the participants" previous experience, the shape the discussion took and the diffractions and interferences that occurred during the process, as well as the priority placed on "being and doing together' and being aware of the (self-)imposed limits while also allowing, even for a short period of time, the dichotomies that characterize and delimit this intercultural encounter to be questioned.

1. Introduction

The kitchen and the very act of cooking have very different social and cultural meanings, depending on their contexts. Indeed, anthropology has always shown interest in this area, though in recent decades it has become integrated with more general studies on food (Faizul, 2018). The kitchen is the place where life happens, is arranged and sustained; it is a political, physical, symbolic and affective space that allows us to reflect on very different themes. Beyond being a physical space, it is a social space "made up of material and symbolic elements, positioned actors, a producer of rhetoric, assumptions, mythologies, contradictions, hierarchies" (Licona García and Cortés, 2019: 172). It is also a space that is of particular interest from a feminist perspective (see, for example, Rosaldo, 1974; Abarca 2006; Gac-Artigas 2009; Williams 2014).

* **Corresponding author / Harremanetan jartzeko:** Mari-Luz Esteban. AFIT-Feminist Anthropology Research Group University of the Basque Country, UPV/EHU. e-mail: ml.esteban@ehu.eus. **Contributed by** Mari-Luz Esteban **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Esteban, Mari-Luz *et al.* «Ethnography of the kitchen: the Women's House, a space for feminist alliance and intercultural encounter», *Ethnography* 2022, Vol. 0(0) 1–18 (https://doi.org/10.1177/14661381211073122)

In this article, we delve into a debate about whether a kitchen was to be installed in a new Women's House in a city in the Basque Country and during which different feminist and cultural approaches emerged. Our unit of observation is the very process that the participants undertook collectively in determining the interior design of the House how the House would be run. There is a traditional saying in Spanish —"hasta la cocina" (lit. "as far as the kitchen")— which takes on three distinct meanings in our study: in its most ethnographic sense, it refers to "going to the core" of said process; in a more literal sense, it means examining the specific discussions about the consequences of there being or not being a kitchen in that space; and in a deeper and more procedural sense, we will address the "kitchen work" itself, including of the reflection and dialogue work that participants undertook collectively to make both this and other decisions with regard to the House itself and how it is run.

The participatory process that we analyze began in 2016 and took place in Hernani, a city of 20,000 inhabitants with a lively social, cultural and political life. It is located about ten kilometers from Donostia-San Sebastián, the capital of Gipuzkoa, which is one of the seven provinces of the Basque Country (located between France and Spain, on the Bay of Biscay and on both sides of the Pyrenees).

Women's Houses are municipal spaces pioneered by the feminist movement, in close collaboration with local institutions. They began to operate in 2003, and since then, they have spread throughout the Basque Country, uniting various types of women's associations, city councils, and equality advisory boards. The Houses run a variety of programs and have legal, labor and sexual health counseling services, and provide targeted support for precarious groups (Esteban *et al.*, 2020). The Women's House in Hernani is called *Kulturarteko Plaza Feminista* (Intercultural Feminist Plaza; in this article we will use the Basque abbreviation, KPF), because it houses both the local feminist and anti-racist movements, the latter of which is led by AMHER, the Multicultural Association of Hernani, a collective that works on issues related to immigration, interculturality and coexistence.

The primary aim of this article is to show how the debate about the kitchen allows us to investigate how association members manage social diversity and the work done to arrive at a consensus. It is a consensus that is, like the KPF project itself, under continuous construction, yet despite being unfinished, it is sufficiently stable to allow different genealogies, histories and practices to intersect and continue nurturing the consensus process. All this takes place in a social context of profound change that creates alliances between different parties and social movements. Achieving such alliances requires an openness to dialogue and to mutual knowledge and reciprocity, and it also enlists the application of specific techniques and know-how. The feminist movement has a know-how that is not always present in other social movements (Authors); it is a know-how rooted in a long history of encounters and disagreements between different feminists, and one that allows for improvisation. It is a dialogue that materializes in the physical and emotional encounter between different people who make up the movement. Such an encounter means that theoretical and political displacements occur, responsibilities are assumed, and the observation of social inequalities becomes more complex. Thus, in using the concept of tension applied in Teresa Del Valle's (2005) study of feminism, the kitchen becomes a metaphor, a juncture of critical and creative tension: it is critical, because it is based on a position of continuous analysis of and judgment about problems and ways of acting; it is creative, because it promotes imagining the possible alternatives and solutions for achieving the necessary consensus.

To achieve our aim, we will first explain our research methodology and then we will describe our ethnographic framework and the social changes that Basque society is experiencing, including the changes related to how politics are done. In the sections that follow, we will delve into the debates and the various views that emerged in the conflict analyzed. We will refer to the dominant cultural assumptions and imaginaries held about cooking in Basque society, as well as to the opposing feminist positions regarding the domestic sphere and cooking, all of which paint a very complex theoret-

ical and practical picture. We will also show that the participatory process is dynamic in two senses: on the one hand, the immediate issues at hand are addressed; on the other hand, as positions become more flexible and are projected into the future, it becomes necessary to build and maintain the conditions that keep relationships from breaking and allow the pact to be renewed and nurtured as many times as necessary. This is a process in which priority is placed on "being and doing together", which entails an awareness of (self-)imposed limits, but also allows for —even if only for a short period of time— the dichotomies that characterize and limit the intercultural encounter to be questioned. We end by discussing the results of our analysis and proposing and discussing some conclusions.

Following Maria Livia Alga (2018), we will assert that the KPF's approach to diversity gives rise to oblique and transversal readings of feminist convictions, which are enhanced by the interferences and diffraction (Haraway, 1999) that occur in the fixed and dichotomous understandings of cultural and gender differences. The result is that the political subjectivities that are formed, despite their being situated in a specific territory and society, tend to transgress and overflow "normative, sexual and cultural, linguistic and geographical borders, which are not identified with a single "world" nor a single category" (Alga, 2018: 147).

2. Methodology

The specific analysis that we present in this article is part of two research projects. The first, "New solidarities, reciprocities and alliances: the emergence of collaborative spaces for political participation and the redefinition of citizenship", is financed by the Spanish Ministry of Economy, Industry and Competitiveness (CSO2017-82903-R, 2018-2021). The second, "Weaving communities from citizen initiatives (2018)", was carried out in 2018 with financial support from the Provincial Council of Gipuzkoa in agreement with the Vice-Rector's Office at the University of the Basque Country (UPV/EHU). The research team is composed of a large group of people belonging to different research groups funded by the Basque government. Specifically, the authors of this article are part of a long-standing research group that specializes in feminist anthropological and sociological studies.

Both projects aim to analyze actions taken by different social movements, focusing on the collaboration, alliances and interactions between different actors and movements (especially feminism, anti-racism, environmentalism and the promotion of the Basque language), because we believe that a close examination of these processes is essential for understanding how the different communities are (re)weaving themselves together. To that end, we implemented a qualitative and ethnographic methodological design, one that combines different research techniques.

The portion of the study that we present here is based on an extended ethnography (two years of fieldwork), where we participated in the process of creating the Women's House in Hernani, the KPF, from the very beginning. In addition, we observed very different events that were related to the House, to feminism and to the dynamics of other social movements in the municipality. Secondly, between 2018 and 2020, we conducted 13 in-depth interviews with people directly involved in the KPF and/or other socio-political initiatives in the municipality. Thirdly, we analyzed a variety of materials: the websites of various collectives and entities, articles from *Kronika* (the local newspaper) and leaflets and manifestos, among others. Lastly, we compared our results with the results corresponding to the study of other Women's Houses in the province of Gipuzkoa, namely the analysis of the processes observed in the same study in the Houses in the cities of Arrasate, Donostia/San Sebastián and Errenteria. Though these Houses are all at different stages of operation, their social outreach and scope are similar; however, they differ in terms of the characteristics of the municipality and the people who participate in them.

3. Promoting social change in a changing society

In recent decades, there have been transformations in all areas of the so-called Western societies, from

the political and economic to the most intimate; changes that have been highlighted further by the Covid-19 crisis. On one hand, we have witnessed the deinstitutionalization of social relations (Touraine, 2005), which have impacted social cohesion and order. Additionally, we are facing an ecological crisis, a care crisis and a civilizational crisis (Herrero, 2016), all of which feminism has denounced in different parts of the world. Likewise, unemployment has spread, working conditions have deteriorated and public services are increasingly precarious, while inequalities between rich and poor have increased (Gaindegia, 2016; Gálvez, 2013). As a result, there is greater pessimism about the role of institutional democracy (Subirats, 2005), and at same time collective responses and different proposals regarding participatory democracy have emerged (Santos, 2004). Indeed, the need to rethink politics has led to an increase in citizen participation initiatives (Martínez-Palacios, 2017). All these changes have influenced how political action and the political subject are conceived, and new forms of political participation have gained importance (Luxán *et al.*, 2014; Esteban, 2015), giving rise to models that are more open, less rigid, and coherent (Esteban, 2015: 83) and processes of subjectivation that are dynamic, contingent and decentralized and made and remade at each step (Berardi, 2013; Diz, 2019; García, 2019).

Basque society is similarly affected by all these new realities. With a population of 3,000,000, the Basque Country has the historic claim of having been an independent nation for a very long time. The region's political and armed conflict of recent decades has had a profound influence, along with all its unfortunate consequences (deaths, torture, attacks, repression by the State, and so on), but the political coordinates have been reconfigured in the wake of the ceasefire and the dissolution of ETA in 2011 and 2018, respectively. Furthermore, throughout most of the Basque Country there are two official languages, Spanish and Basque (Euskara), though they are far from equal; additional languages are also spoken in the territory as a result of migration. Coexisting in different languages and simultaneously engaging in defense of the Basque language entails a great deal of social activity.

In this context, feminism is a growing social movement, becoming a transversal axis and a significant impetus for many other initiatives and movements. Feminists are proposing more dynamic and horizontal participation methods, paying attention to relationships between people and taking care of collective processes (Esteban *et al.*, 2020). Furthermore, feminist debates are becoming increasingly complex, refining the analysis of social and gender inequalities through an especially interesting intersection between queer, transfeminist, anti-racist and class perspectives and positions favorable to the recognition of the people's sovereignty. Likewise, such debates try to address the system of privileges and the lack of material and symbolic redistribution among women and the population in general from an approach that is accountable and self-critical. In fact, one of the challenges that the Basque feminist movement has on its political agenda and which it laid out at the 5th edition of Euskal Herria Feminist Conference (Durango, Bizkaia, November 2019) is to commit to a practice that is anti-racist, intersectional and decolonial. It is precisely for this reason that we believe that a space shared by feminist and anti-racist groups, such as the KPF, is a privileged laboratory not only for analysis, but also for learning and experimenting with necessary social transformations.

4. The participatory process at the House and the debate over the kitchen

Before focusing on the debate at KPF, we would like to make two general points about Women's Houses in the Basque Country. This first is to note that the horizontality of the dynamics surrounding House management is a key principle of those very dynamics, but this does not mean that there are no internal power imbalances. The second is that the Women's Houses are jointly run by the feminist movement and local institutions, where many times the institutions themselves initiate the establishment of a House in response to community demands, which in many cases are quasi-historical. This is the case for the KPF in Hernani.

Hernani's Women's House project emerged from a proposal by the city's Equal Opportunity Board (a body made up of citizens, equality specialists and political party representatives), following a participatory process that began in 2016. This process, which is still ongoing, was originated by the city council, but facilitation has been provided by a cooperative that is specialized in participatory processes and group-facilitation methodologies. Members of the relevant associations and groups, as well as individuals, participated in the process, and the meetings served as the primary space for debate and decision-making, although there was also a steering group that coordinated and led the process.

As we have already noted, the KPF will be home to various groups and individuals involved in the city's feminist movement as well as the immigrant association AMHER. The members of AMHER represent more than twenty different countries and the association has various working groups. One such group is made up of women, and it participates in local feminist initiatives, including the creation of the Women's House. Thus, in addition to collaborating, the two movements intersect. This enriches the process, and it also blurs, to a certain extent, the boundaries between the two groups. All those participating in the House process view the KPF as a point of reference against all types of discrimination, although they are also aware of the difficulties involved in managing "diversity" in its most general sense.

The people involved in these two groups are also involved in other initiatives in the city: youth movements, political parties and unions, environmental and cultural associations, and collectives that support the Basque language or the LGTBI community. Furthermore, many of the same people are engaged in more than one initiative at any one time, a phenomenon that characterizes Basque activism in general, as it tends to be multiple, multi-sited and interrelated. The fact that people in a city like Hernani participate in multiple initiatives, know each other, come together and collaborate enriches the project by involving a multiplicity of perspectives and understandings of sources of oppression. This guarantees a more inclusive vision regarding the building the common space, one that is supported by the affective relationships that bind the participants together. All of this directly affects their desire for "being and doing things together", an idea inspired by various authors (Gil, 2011; Kypriotaki, 2012; Esteban *et al.*, 2020). As we will explain later, this desire is an effective way to weave bridges between different groups and create solidarity networks.

Some of the most intense discussions that emerged in our analysis of the participatory process were related to the physical and architectural design of the house, the most illustrative example of this being the decision about whether to dedicate a specific space to the kitchen. This topic came up in all the conversations we had with the participants without us needing to prompt them. Some people gave it more importance than others, but it was a recurring topic, and everyone had an opinion about it. Almost from the beginning we realized that there was a sticking point there that would allow us to discern the process as a whole.

At the time when the interior layout of the House was being decided, some women from AMHER proposed that space be set aside for a large kitchen and a day-care for children, sparking a very passionate debate around both issues. Although the issues are related, we will save the second for another occasion and focus on the controversy over the kitchen, which was more extensive.

Aisha, a woman who is very active in AMHER and also has ties with the feminists, summed it up by saying that the kitchen is a fundamental space for the women in AMHER. She argued that the space that they were using at that time, where they would meet and have their Spanish classes, was also equipped for cooking. It was a space that they used often, not only because cooking and eating as a group was the main event of any meeting. But it was also the case that having a space to prepare food enabled some women to earn money, because they received orders for "food from different countries around the world." Ángel, who is of Latin American origin and a member of the AMHER collective,

noted in a conversation about women from African and Latin American countries: "Women from Morocco or other African countries have the custom of inviting people to their home and receiving them in the kitchen, and they always offer you something to eat as a way of expressing that you are welcome."

But for some other women at the meeting, the idea of the kitchen generated great contradictions, and an extremely intense debate broke out. These women, who had attained a high degree of education, had spent many years in the feminist movement and/or had been union members, and were experienced in debating political topics (in both Spanish and Basque), were radically opposed since they problematized the fact of relating the specific spaces for women with the domestic tasks traditionally assigned to them. Additionally, they thought that using the KPF premises as a place of employment for some would be a very difficult issue to administer and there would be endless consequences.

Mari Karmen, one of the women opposed to the kitchen, reported that what had caught her attention most was the reaction of the young feminists who were not members of AMHER. Not only did the young women not understand the debate, they thought that having a kitchen would be a good opportunity to prepare and eat vegan food together, bringing the topic of food into the idea of group mutual care. In later conversations with her, she added that over time she had realized that what these young women stood for at the time was becoming the general trend in some feminist or mixed associations. These were very young women, many of them either high school or university students who participated in their schools' feminist groups as well as in the Urumea transfeminist squat coordinated the young Basque feminists who organize *Udaleku Feministak*, or feminist summer camps. As antispeciesism and veganism are found in the latter two spaces, any activity relative to the kitchen has become a political issue.

As we have noted, the deliberations, which took place over several sessions, were complicated and of interest to everyone, although the idea of imposing some kind of limit on the physical space of the kitchen prevailed. The final agreement was that the House would have simple (rather than industrial grade) equipment for cooking, but the space would be multipurpose, meaning that meetings and other types of activities could be held there.

Before going further into the details of the debate over the kitchen, let us first review different feminist readings on the kitchen, readings that, as we will see, are reflected in the various positions found in the KPF. We also take into account the cultural significance that the act of cooking has acquired in the Basque Country in recent years, as we believe that this significance undoubtedly influences the feminist position "against the kitchen". We start with the latter.

4.1. Feminist readings of the kitchen and the domestic

Mabel Gracia-Arnaiz (2014: 26), following the work of George Peter Murdock and Caterina Provost (1973), points out that

> *"ethnographically and historically, women have been and are —with the exception of those who are part of elite groups in differentiated societies— the people responsible for daily sustenance, especially in relation to the tasks of provisioning and preparing family meals."*

Gracia-Arnaiz also points to Stephen Mennell (1985): "In societies where a gender-differentiated kitchen exists, the role of the cook —the *chef*— is male" (Gracia-Arnaiz, 2014). This distinction between cooks and chefs is present in very different societies, including Basque society. The differentiation is based on a gender-based division of labor that does not view many of the tasks performed by women to be labor; this division, in turn, is articulated, though not always linearly, through the dual characterization of space and a differential allocation of prestige. Thus, everything related to the female world is considered "domestic" and less prestigious, and everything related to the male world is considered public and more prestigious.

Gathering to eat is fundamental to the social imaginary and identity in Basque culture. It is an act that tends to be related to consumption and leisure,

and one that we cannot fail to link with the primordial material and symbolic place occupied by what are known as gastronomic societies (*txokoak* in Basque), which are member-operated clubs for private recreation and gatherings. Given that these societies are present in cities and neighborhoods, they play an important role in socialization, social engagement and the creation of networks of influence and power. But gastronomic societies have traditionally been led and run by men, and women have been socially excluded. Today, women can be members in the vast majority of cases, but they are not always allowed in all the spaces, especially the kitchen (Farapi, 2010).

The most paradigmatic example of the transformations that have occurred around the kitchen in the Basque Country is represented by the male cooks who run the Michelin-starred restaurants that are the drivers of what is known as new Basque cuisine. It is a highly prestigious profession and very clearly gendered, and since the creation of the Basque Culinary Center (which is part of the Mondragon University, an affiliate of the Mondragon Corporation) it is a profession that is now associated with a university degree. Once again, we see the separation between female cooks and chefs, a phenomenon criticized among feminists; one such criticism comes from anthropologist Del Valle (2000), who has described this male-dominated professionalization of the kitchen as a usurpation of women's knowledge: "… a usurpation that implies the denial of genealogies despite the fact that they make references to their grandmothers to highlight the traditional nature of their stews" (2000: 55).

The gender-differentiated kitchen, and the specific transformations and divisions engendered by this division in Basque society, allow us to understand the "disaffection" noted by feminists in our study who are local-born and over the age of 40 and the disagreement over whether to give the kitchen a central space in the KPF. Those feminists are also influenced by a feminist tradition that is critical of identifying women with the domestic arena, as we will detail below; this critique extends to the name given to this particular Women's House, *Plaza Feminista Intercultural*, where *plaza* (a public space) was expressly chosen over Spanish *casa* (house) or Basque *txoko* (private spaces).

In the second half of the 20th century, feminists belonging to the hegemonic tradition of the time, that is, those in Anglo-European societies, began to problematize the link between the traditional role of women and the domestic sphere. The home was primarily seen as representing a symbolic space where the discipline and oppression of women occur. In the words of Priscilla Gac-Artigas (2009: 512), "everything belonging to the intimate (and exclusive) sphere of the woman, the family or the home was rejected because it was considered to be the cause of the subordinate status of women in a patriarchal society". This movement, with its desire to "integrate" women into the social sphere and its belief that "the personal is political", politicized everything that happened in the intimate sphere, in the home and in social relations, among other arenas. Symbolically, we could say that this breaks with the ideology and archetype of a woman and a perfect "housewife".

This approach was also supported by the work of authors such as Michelle Rosaldo (1974), who, like other contemporary feminist anthropologists, investigated the symbolic causes of the subordination of women. She showed how the conceptualization and opposition between the "domestic" and the "public" arose at one point in Western history and provided the basis for a structural model that allows the subordination of women to be identified and explored, arguing that men have acquired authority, hierarchy and rank through their actions in a separate political world (Maquieira, 2001). However, Rosaldo (1980, 1983) soon revised her theory, aware of the universality of the categories and theories used and of the essentialist dualist schemes. In addition, it became clear that the public/private dichotomy has a clear ethnocentric bias, and that it cannot be applied as an absolute model of analysis in the West, either, due to the difficulty of defining the limits and character of these areas as well as the complexity of reality (Esteban and Díez, 1999; Maquieira, 2001).

The feminists who problematized the kitchen as a feminized space described it as "rejecting the

housewife role and the actions that accompany it, while focusing on the attempts to integrate women into previously male-dominated public spaces" (Williams, 2014, 2016). But as Stacey J. Williams (*ibid*) points out, although they have been less frequently discussed, during that period there were also proposals that suggested engaging with cooking in ways that were subversive and challenged patriarchal institutions. Some feminists suggested time- and labor-saving cooking methods, encouraged men to cook, and proposed that women make money from cooking. These ways of politicizing cooking "were meant to increase women's control of economic resources", bringing about "a more gender-equal world" (Williams, 2016: 270). Moreover, there are other studies that show how the kitchen, as a collective space, can be a space for women's self-care and collective care. Meredith Abarca (2006), for example, in her work on views of food and the world from working-class Mexican and Mexican American women, showed how cooking allowed them to express themselves, strengthen family relationships, and create a world of shared meanings with other women. Thus, Abarca (*ibid*), explores the importance of the knowledge found in the practical, concrete, and temporal aspects of the ordinary practice of everyday cooking. In this regard, the impact of certain novels by renowned Latin American writers is also relevant. Gac-Artigas (2009) analyzed the work of Rosario Castellanos, Isabel Allende, Laura Esquivel and Rosario Ferré, whose works turn the kitchen into a space for women's self-discovery and liberation.

4.2. *"Entangled in the kitchen": the collective management of diversity and consensus building*

Returning to the KPF, we observed that for some immigrant feminists, cooking had both a practical and cultural value; additionally, feminists of a certain age were, to one extent or other, "feminists born and trained to be against cooking", and the youngest Basque women did not see where the problem lay. Many of these young women adhere to a kind of feminism that, inspired by approaches that emerged in recent decades, turns over many of the previous theories, and they are also influenced by a combination of alternative perspectives: from ecofeminist approaches and spiritualist worldviews, to the growing influx of communal feminisms and the postcolonial and decolonial theories of Latin American thinkers and activists. For them, the motto "put life at the center", which has become one of the signals of feminist identity today, allows them to fully accommodate the idea that cooking as a group is positive and can even be transformative. As already noted, these young feminists participate in other political spaces that are committed to vegetarianism or veganism, where cooking and eating as a group is directly linked to one's politics.

Capitalist culture accelerates, commodifies and reduces time and space in the kitchen and, more generally, the daily tasks that sustain life, and many feminists are very sensitive to this process and demand time and space for collective care, turning it into an anti-capitalist symbol. And as we noted above, the decolonial critique has led many women to broaden, question and revise ideas and practices about intersectionality, geopolitics, social class, care, and the politicization of the personal. In the Basque context and throughout the Spanish State, this shift has come from the knowledge of and questioning by immigrant feminists who define themselves as racialized.

Meanwhile, what has happened or is happening with the kitchen in other Women's Houses in the Basque Country? Broadly speaking, in many of them there is a space equipped for preparing coffee or tea, or to heat food, but at the same time there is a tendency to put limits on this space in some way, sometimes intentionally and other times not. And if we leave our borders and focus on other spaces, such as the Women's House known as the "Centro Interculturale delle donne di Ramia" in Verona (Italy), an intercultural center for women that we are very familiar with, it allows us to find other nuances. The House is included in the social services provided by Verona's City Council and its operation is inspired by ideas from difference feminism. They conceive of that House as a new space, a "third space", where the kitchen is a multipurpose space and a meeting place, serving as a space to eat together and, above all, a space that

promotes the feeling of "being at home"; it is also is an economic strategy for people who have fewer resources. In addition, the House places great importance on recognizing all kinds of traditional and generational knowledge that is left out of the market. Nevertheless, it's worth noting that in this case there was also a discussion about the suitability of using the space to carry out economic activity, the result being the creation of a cooperative.

Returning to the KPF, an aspect that we want to highlight from the debate about the kitchen is that it made many of the participants understand that what for some symbolized the danger of engaging in gender essentialisms was for others a kind emancipation —a practical resource for their economic survival and gathering with others. Being able to listen to each other and bear witness to power relationships and privileges implies a willingness to move and go beyond one's principles, or at least to make them more flexible. Although in the end the participants opted to create the kitchen, it is viewed as a space for a multitude of uses. In short, thanks to the participants' previous experience and the facilitator's help, they were able to identify disagreements, leave room for dissent, and build consensus among everyone. This is reflected in the way meetings ended, with everyone sitting in a circle and facing the other participants; once the meeting was over, lively conversation followed, both in the meeting room and on the way home.

Feminism has ample expertise in the above regard, and the kitchen itself was used as a metaphor during the process. It is evident that this new House is already generating physical gatherings and will generate more in the future. Political and emotional relationships, especially when the politics of intimacy occur within them (Ahmed, 2004), are embodied and lead to the reinvention of ways to do politics (Guilló-Arakistain, Esteban and Luxán-Serrano, 2020). In this regard, the key elements were how the sessions ran —always in movement and maintaining physical contact, alternating small group work with work in the larger circle— and the facilitators' ability to give voice to all voices, soothe heated spirits and redirect the discussion when necessary. However, it bears repeating that managing diversity is not without complications and interferences. But it is precisely these interferences, as we will comment on in the next section, which allow for the development of a self-critical and regenerative approach. Thus, diversity is not merely an objective; rather, it is more than anything else an exercise in unlearning certain attitudes, questioning one's own view of things and making the journey together.

The Women's Houses are spaces where new forms of solidarity, new methods and new ways of doing politics are being tested; spaces where "community is made", a community rooted in and committed to specific political, social and cultural coordinates, while being aware of the need for thinking that goes beyond geographical and human borders. They are laboratories where horizontal forms of learning, mutual knowledge, conflict management, and practices of care regarding process, concrete projects, and the group are tested, experienced and developed. This is true even when projects sometimes fail. Because what is important is not the final product, but the path traveled together.

5. The kitchen as a breaking of dichotomies and the renewal of politics

The debate analyzed here also allows us to highlight the importance of women's participation in urban planning and in all decisions related to the projects in which they are involved. This is true especially when these projects emerge from the joining of institutions and social movements, given the risks that are involved. In this sense, the real decision-making capacity that the participants have had in some Women's Houses has been quite a controversial issue (Esteban *et al.*, 2020), not only due to the very processes involved in their creation but also due to the social limitations placed on the ability of certain groups, for example, immigrants, to participate in political decisions.

From the beginning of the process, both the feminists and the city government made the effort to bring together people of different origins, social positions and ages. In general, most of the people

interviewed have been satisfied with the extent to which women have participated in the process and the diversity among the women. Nevertheless, with the help of the participants, we have identified a set of obstacles that particularly affect immigrant women. Such obstacles are not always easy to deal with, and in some cases, they are not even readily apparent. Obstacles include their lack of time for participating in socio-political action due to their employment situations and/or the lack of childcare networks; city policies that support multiculturalism (which is also promoted by progressive groups) but which often tend toward folklorism (particularly the privileging of activities related to food); and, most crucially, the division between "us" and "them", which is present even in the mind of the youngest women. Even though this issue was always questioned when theorized about in the interviews, in more spontaneous speech and discussions, this issue often emerged.

It is not our intention to present a simplistic or excessively positive reading of what happened, nor to present the case analyzed as a fluid and linear process, running cleanly from back to front and always moving forward. If we probe further into the pitfalls of the process, including the elements discussed in the previous paragraph, we find certain key points that conditioned the debate and that will still be present in the future. These key points include differences in education and experience in politics among the women in the KPF and the format in which the discussion took place, which has some features that favor mutual understanding and others that do not.

When Aintzane, the group facilitator, describes her experience with the process, several essential elements appear. On the one hand, the women who were strongly opposed to the kitchen had a higher level of education and greater linguistic competence from their years of being trained to defend their position in a group discussion. But this was not the case for many others —whether they were in favor of the kitchen or they did not understand what was happening— who merely asked, time and again, why not have a kitchen. On the other hand, there is the fact that the sessions were held in Basque, facilitated by an interpreting service that was funded by the city council. While all of the women accepted the bilingual nature of the deliberations, this meant that the debate became less fluid at the most heated moments. This shows the difficulty that can arise when different languages are involved, even when technical and economic resources are available. In any case, the role of the facilitator was essential (and praised by all); even when she did not fully understand what was happening, she made the effort to ensure that the floor was held equitably and to soften and streamline the tone of the discussion. Aintzane uses the term "orthopedic" to refer to communication during the discussions: arguments in favor of the kitchen made in Spanish, often expressed in a less than fluid manner, and arguments against the kitchen made in Basque, often dynamically and loudly; words that cross each other but don't make it to the other side. In our interview with Pilar, she elaborates on the simultaneous interpretation used to facilitate communication between everyone and adds an arresting visual image of these moments: "Yes, but the immigrant women really stood out, they wore little antennas (from the headsets they wore to hear the interpreter) and sat together".

From our conversations with many of the participants, we have concluded that they were all aware of, or at least intuited, all the factors mentioned here —despite their not being made explicit as such— and in the end this awareness had a direct effect on participants' capacity to compromise and take more flexible positions. In other words, both the participants and the facilitator looked for mechanisms that could compensate for, even if only partially, the unequal position that some of the participants found themselves in. All of this is in keeping with their feminist philosophy and their long-standing political tradition.

But we would like to take the analysis a little further and delve into a couple of aspects mentioned in the introduction, aspects that strike us as defining. The first has to do with the concrete form in which the discussion took place, which requires additional detail about how the meetings were held. From time to time, those participating in the House cre-

ation process would meet to discuss various topics, and everyone would sit in a large circle. A circular arrangement "composed of a multiplicity of voices and hands (…) generates a specific way of sharing knowledge" (Cima, 2020: XIV), favors eye contact, listening and paying attention to others, as well as a sense of group belonging; this, in turn, enhances the space, making it more welcoming and promoting reciprocity (Cima, 2020). The center is an empty but non-neutral space which symbolizes, according to Alga (2019), the encounter and the possibility of thinking without predetermined schema. The facilitator is also in the center, occupying this special place while also being perfectly aware of her position. She moves around and "appears and disappears", synthesizing what is being said, asking questions—sewing the stitches that make it possible to baste the difficult discussions together. And all of this takes place within a framework of attachments, which play a key role in collective action. The attachments are learning and dialogue: the presence, the encounter between bodies that open themselves to relationships and to different languages and knowledge.

The second aspect, which is related to the previous one, has to do with a comment by Mari Karmen in which she stressed, in a tone that evoked the emotion she felt at that moment, that she was genuinely astonished by the reaction of some of the young feminists, who, despite having been schooled in Basque feminism, were surprised by what was happening. From our point of view, it was an instance of what Donna Haraway (1999) would call a diffraction, an interference in individual and collective thought. This optical metaphor allows Haraway to include two aspects that she considers key to the critical exercise: the ability to look from the other side and to recuperate views that have been kept outside the hegemonic rationales. This is precisely what we believe was happening in this scene and in many other similar scenes that occurred during the process of creating the KPF.

This would lead us to conclude that a map of diffractions and interferences that is drawn in a political and emotional territory that is ripe for criticism and self-criticism, such as the KPF, can introduce fissures and raise meaningful questions; in this case, we refer to the questioning of the dichotomous perspective of "us/them", "native-born/immigrant feminists". In other words, the accumulation of interferences facilitates an oblique, transversal look at feminist thought and action, which fractures verticality and the antagonistic gaze, and disposes those involved toward "widening the circle" (Alga and Cima, 2020) that they construct together. As long as circumstances are favorable, that is, as was the case here.

In other words, in essence we believe the kitchen debate served to break, regardless of whether it was temporarily so, the dichotomy "us = native-born population/them = immigrant population." This dichotomy, despite the anti-racist ideology of social movements, is not easy to overcome, and feminists like Itsaso, another of the participants in the House creation process, consider it crucial that participants be very aware of the intersection of different factors and, crucially, not forget the importance of racialization, social class and educational training.

Recall that the process has not ended, not only because the building itself is not yet ready and the internal operating protocol is pending, but also because the work on feminist viewpoints continues. In other words, it is a consensus still under construction, not because the decisions made are not firm, but because it can be reviewed and completed later, and the shapes drawn do not have to be linear. In this regard, it is interesting that some of the participants who had not fully understood the reactions "against cooking" later told us that, by taking part in other activities in the city, they were able to broaden their perspective. Specifically, they cite a conference held in April 2018 as a tribute to Empar Pineda, a long-established Spanish feminist, who had been born in Hernani, as a watershed moment. The conference's organizers made an effort to integrate the feminist genealogy of the last five decades, which made it possible to contextualize the feminist proposals and analyses from the 1970s and 1980s.

In our fieldwork, we have noticed that alliances emerge along with a renewed way of working together, which we have called "being and doing things together". This also happened in the case

of Hernani. In this "being together", the projects and platforms that are made up different groups and created on the fly through the participation of everyone are of the utmost importance. In general, we have also found that initiatives of a limited duration and that require a temporary commitment are particularly successful in today's social movements. This change is leading us to rethink socio-political participation. Activists involved in different movements have linked "being together" and "doing things together" with a renewed model of understanding citizenship that is based on active participation and clashes with other traditional ways of understanding citizenship, which are based on merely administrative or legal criteria.

The KPF's kitchen has been redefined as a privileged feminist space in which to think about all these questions. It is not because it is the natural space for being a woman or because women's relationship with the kitchen has prevailed, nor is it because it is believed that having a kitchen and cooking is better for a political space. Rather, it is precisely because this kitchen symbolizes the debate, the process, the listening, and the agreement reached among various political subjectivities that reformed themselves through the process. Although all those involved remember the deliberations as having been difficult, these debates are now part of the *body's memory* (Del Valle, 1997), which can be evoked and reactivated at another time as a way hold on to the awareness of how arduous but necessary it is to have a policy that is aware of intersectionality and internal and external inequalities, which promotes thinking that is constantly moving. Understanding a social action as a physical and emotional phenomenon provides the appropriate framework for investigating the place that community, relationships and emotions have within them (Guilló, 2020). The individual and collective subjectivities that we refer to in this article are in continuous transformation and allow the emergence of other ways of doing feminist politics and living a feminist life together (Ahmed, 2018).

Similarly, the specific case of the kitchen illustrates quite well the relevance of anthropological work. Being able to know the details of the discussions that took place, observing them in situ, and relating them to the feminist and cultural viewpoints that are behind ideas and experiences helps us better understand the limits and the complexity of the policies and social changes that are occurring and/or being proposed and be able to render account. Hernani's KPF is a project still under construction and one that will continue to be under construction after it begins operating. For only when under construction is it possible to face the dilemmas and difficulties of a politics that is aware of its (own) limits but aims for new agreements and consensus.

6. References

Abarca M 2006 *Voices in the Kitchen: Views of Food and the World from Working-Class Mexican and Mexican American Women*. College Station: Texas A&M University Press.

Ahmed S 2004 *The Cultural Politics of Emotion*. New York/Abingdon: Routledge.

Ahmed S 2017 *Living a Feminist Life*. Durham (US): Duke University Press.

Alga ML 2018 *Etnografía "terrona" de sujetos excéntricos*. Bellaterra: Barcelona.

Alga ML 2019. Construyendo en común la investigación. El trabajo con los movimientos sociales. In: *Tejiendo comunidades desde iniciativas ciudadanas* (Summer Courses UPV/EHU), Donostia, Spain 17-18 June 2019.

Alga ML and Cima R (eds) 2020 *Allargare il cerchio. Pratiche per una comune umanità*. MeTis International Journal. Bari: Progedit.

Berardi (Bifo) F 2013. *La sublevación*. Barcelona: Artefakte.

Cima R 2020 La circolarità dei saperi. In: Alga ML and Cima R (eds) *Allargare il cerchio. Pratiche per una comune umanità*. Bari: Progedit, pp. XIII-XVII.

Curiel C 2019 Mujeres en la cocina de la mayordomía. Prestigio y costumbre en la mixteca de Oaxaca. *Entre diversidades* 6, 1(12): 163-194.

Del Valle T 1997 La memoria del cuerpo. *Arenal-Revista de Historia de las Mujeres* 4(1): 59-74.

Del Valle, Teresa. 2000. La organización del tiempo y del espacio. Análisis feminista de la ciudad. *Zainak* 19: 53-60.

Del Valle T 2005 El potencial de la tensión y su aportación a la antropología desde la crítica la feminista: fuentes, procesos y tipologías. In Maquieira V *et al.* (eds) *Democracia, feminismo y universidad en el siglo XXI*. Madrid: Instituto Universitario de Estudios de la Mujer/Universidad Autónoma de Madrid, pp. 227-242.

Diz C 2019 Maneras de vivir: emoción, política e identidades en movimiento. *ANDULI, Revista Andaluza de Ciencias Sociales* 18: 93-117.

Esteban ML 2015 La reformulación de la política, el activismo y la etnografía. Esbozo de una antropología somática y vulnerable. *Ankulegi* 19: 75-93.

Esteban ML 2019 *El feminismo y las transformaciones en la política*. Barcelona: Edicions Bellaterra.

Esteban ML and Díez C 1999 Introducción. Antropología Feminista: desafíos teóricos y metodológicos. *Ankulegi-Revista de Antropología Social* 4: 9-28.

Esteban ML *et al.* 2020 *Tejiendo comunidades desde iniciativas populares. Resumen del Informe de investigación*. Bilbao: Servicio editorial de la Universidad del País Vasco (UPV/EHU).

Farapi/Consultoría de Antropología Aplicada 2010 *La relevancia e impacto social de las sociedades gastronómicas*. Vitoria-Gasteiz: Defensoría para la Igualdad de Mujeres y Hombres/Eusko Jaurlaritza.

Gac-Artigas, P 2009 La cocina: de cerrado espacio de servidumbre a abierto espacio de creación. *Destiempos. Revista de Curiosidad Cultural* 19: 512-522.

Gaindegia 2016 *Des-herria. Hausturak eta etenak Euskal Herrian*. Andoain: Gaindegia/Ipar Hegoa.

Gálvez L 2013 Una lectura feminista del austericidio. *Revista de Economía Crítica* 15 (1): 80-110.

García E 2019 *La construcción social del activismo en Madrid durante el ciclo 15M: subjetividades políticas y resistencia antiausteritaria*. PhD Thesis, Universidad Autónoma de Madrid, Spain.

Gil S 2011 *Nuevos feminismos. Sentidos comunes en la dispersión*. Madrid: Traficantes de Sueños.

González I (1999) La dimensión social de la cocina desde la antropología de la alimentación. In: *Alimentación y cultura: actas del congreso internacional, 1998, Museo Nacional de Antropología, España*. Huesca: La Val de Onsera, pp. 245-257.

Gracia-Arnaiz M 2014 Alimentación, trabajo y género. De cocinas, cocineras y otros trabajos domésticos. *Panorama Social* 19: 25-36.

Guilló-Arakistain M 2020 *Hilekoaren politika eta kultura alternatiboen etnografia bat: genero-konfigurazioak, gorputz-ahalduntzea eta ezagutza kolektiboak*. PhD Thesis, University of the Basque Country, Spain.

Guilló-Arakistain M, Esteban ML and Luxán-Serrano M 2020 Le case delle donne nei Paesi Baschi: dibattiti, processi e alleanze. In: Alga ML and Cima R (eds) *Allargare il cerchio. Pratiche per una comune umanità*. Bari: Progedit, pp. 25-41.

Haraway D 1999 La promesa de los monstruos: Una política regeneradora para otros inapropiados/bles. *Política y Sociedad* 30: 121-163.

Herrero Y 2016 Economía feminista y economía ecológica, el diálogo necesario y urgente. *Revista de economía crítica* 22: 144-161.

Holden TJM 2013 The overcooked and the underdone: Masculinities in Japanese food programming. *Food and Culture* 13 (1-2): 39-65.

Hupkens CLH 2000 Social class differences in food consumption. The explanatory value of permissiveness and health and cost considerations. *The European Journal of Public Health* 10 (2): 108-113.

Ibrahim FH 2018 Kitchen Anthropology: Understanding Food, Cooking and Eating in Bruneian Middle-Class Families. Institute of Asian Studies. Working Paper Series 38.

Kypriotaki AO (2012) Una nueva zona temporal del ser (los griegos y la nueva psicodelia). Available at: https://blogs.publico.es/fueradelugar/1625/%c2%bfy-si-no-hiciesemos-nada (accesed 30 October 2020).

Licona E, García IC and Cortés A 2019 El espacio culinario. Una propuesta de análisis desde la Antropología de la Alimentación. *Antropología Experimental* 19 (15): 165-172.

Luxán-Serrano M *et al.* 2014 Metamilitantzia. Herri-mugimenduen baitatik gogoetak. *Jakin*, 203: 93-107.

Maquieira V 2001 Género, diferencia y desigualdad. In: Beltrán E and Maquiera V (eds) *Feminismos. Debates teóricos contemporáneos*. Madrid: Alianza Editorial, pp. 127-190.

Martínez-Palacios J (ed) 2017 *Participar desde los feminismos. Ausencias, expulsiones y resistencias*. Barcelona: Icaria.

Mennell S 1985 *All Manners of Food. Eating and Taste in England and France from the Middle Ages to the Present.* London: Basil.

Moore HL 1991 *Antropología y feminismo.* Madrid: Cátedra.

Murcott A 1983 Cooking and the cooked: A note on the domestic preparation of meals. In: Murcott A (ed) *The Sociology of Food and Eating*. Aldershot: Gower, pp. 78-85.

Murdock GP and Provost C 1973 Factors in the division of labor by sex: A cross-cultural analysis. *Ethnology* 2 (2): 203-225.

Ortner S 2006 *Anthropology and Social Theory culture, power and the acting subject.* Durham: Duke University Press.

Ortner S and Whitehead H 1996 Indagaciones acerca de los significados sexuales. In: Lamas M (ed) *El género. La construcción cultural de la diferencia sexual.* México DF: PUEG, pp. 127-180.

Rosaldo M 1974 Mujer, cultura y sociedad: una visión teórica. In: Harris O and Young K (eds) *Antropología y feminismo.* Barcelona: Anagrama, pp. 153-180.

Rosaldo M 1980 Use and Abuse of Anthropology: reflections on Feminism and Cross cultural Understanding. *Signs* 5 (3): 389-417.

Rosaldo M 1983 Moral/Analytic Dilemmas Posed by the Intersection of Feminism and Social Science. In: Haan N *et al.* (eds) *Social Science as Moral Inquiry*. New York: Columbia University Press, pp. 76-95.

Rosler M 1975 *Semiotics of the Kitchen (videoperformance)*. New York: Courtesy Electronic Arts Intermix (EAI).

Santos BS 2004 *Democratizar la democracia. Los caminos de la democracia participativa*. México DF: FCE.

Subirats J 2005 Democracia, participación y transformación social. *Polis Revista Latinoamericana*. Avalaible at: *http://journals.openedition.org/polis/5599* (accesed 17 September 2020).

Touraine A 2005 *Un nuevo paradigma para comprender el mundo de hoy.* Barcelona: Paidós.

Williams SJ 2016 Subversive Cooking in Liberal Feminism, 1963-1985. *Gender and Food: From Production to Consumption and After*. Bingley: Emerald Group Publishing Limited, pp. 265-286.

Williams SJ 2014 Feminist Guide to Cooking. *Contexts* 13 (3): 59-61.

Korrika ekitaldian parte hartzeak dituen onurak

Pertsona asko biltzen direnean, helburu jakin batekin, gizartearen kohesioa eta pertzepzio positiboak bultzatzen dira. Kasu horien ondorio psikologiko positiboetan sakontzeko, euskararen aldeko Korrika aztertu dute adituek. Parte hartzeak «eferbeszentzia kolektiboa» deritzon fenomenoa dakar, *Kama muta* izenekoa, esperientzia partekatuek bultzatutako emozio bizia. Horregatik, kontuan hartu behar da horrelako ekitaldiak egitea osasungarria dela parte-hartzaileentzat.

Sinkronia emozional horrek batasun, lotura eta kontzientzia kolektiboaren sentimenduak sustatzen ditu. Beraz, kohesio soziala, konfiantza eta ahalduntzea handiak dira. Hala ere, ekitaldi horretan parte hartzeak epe luzera izango dituen ondorioak esplorazio-eremu bat dira oraindik. Ikerketa batzuen arabera, ondorio positibo horiek aste batzuetan iraun lezakete. Azterlan honen helburua da *Korrika*n parte hartzearen ondorioak aztertzea, hau da, ekitaldian zehar izan den dinamika emozionala eta norbanakoaren eta taldearen ongizatean izan duen eragin iraunkorra.

Korrika erritu kolektibo bat da, euskal hizkuntza eta kultura zaindu eta ospatzearen sinboloa dena. Horregatik, oso aukera ona eskaintzen du prozesu kolektiboak aztertzeko. Soziologia akademikoaren testuinguruan, azterlanak prozesu kolektiboen eredu neodurkheimndarra berresten du, hau da, gizartearen antolamenduaren jarduerek giza emozioetan duten eragina aztertzen duen teoria. Kasu honi aplikatuta, teoria horrek erakusten du *Korrika* ekitaldian parte hartzeak gizartearen kohesioa hobetzen duela, gizartearekiko konfiantza indartzen duela eta gizabanakoak eta taldeak sendotzen duela. Zehazki, erritualarekin bat egiteak gizarteratze-sentimenduak pizten ditu, eta eragin positiboak ditu gizartearen onarpenean eta eguneratzean.

Kama muta

Parte-hartzeak norberaren jabekuntzarako eta gizartea hobetzeko sentimendurako ere laguntzen du. Pozak eta *kama muta*ren sentsazioak (lotura emozional bizia) funtsezko garrantzia dute ekitaldi kolektibo horietan. Eta, denborarekin, erritual kolektiboen partaide izatearen, onartzearen eta gizartean konfiantza izatearen sentimenduetan dituen ondorio positiboek iraun egiten dute, osasun psikologikoa sustatuta. Beste gertaera askori lotutako berehalako emozioak desegin daitezkeen arren, badirudi eferbeszentzia kolektiboaren efektu zabalenek inpaktu iraunkorrak dituztela.

Etorkizuneko ikerketek dinamika horiek aztertzen jarraitu behar dute, eta azterketaren mugak aztertu, hala nola kontrol-taldeen beharra eta neurtzeko tresna osatuagoen potentziala.

«Erritual kolektiboen ondorio positiboek iraun egiten dute denborarekin»

Emaitzak argiak dira. Hasteko, *Korrika*n parte hartzeak ondorio psikologiko positiboak ditu gizarte-ongizatean. Bigarrenik, ondorioztatzen da eferbeszentzia kolektiboa izan daitekeela parte-hartze kolektiboaren ondorioen mekanismo nagusietako bat, eta bereziki garrantzitsua dela hura mantentzeko. Hirugarrenik, *kama muta* hori emozio garrantzitsu gisa sortzen da, eta arreta jarri behar zaio etorkizuneko ikerketetan. Eta, azkenik, bilera kolektiboak antolatzea eta parte-hartze kolektiboa osasuneko prebentzio-prometan baloratu beharreko faktoreak dira.

InpAkta 00 (2023) 203-223

https://doi.org/10.3389/fpsyg.2023.1095763

Korrika, running in collective effervescence through the Basque Country: a model of collective processes and their positive psychological effects

Jon Zabala[1*], Susana Conejero[1], Aitziber Pascual[1], Larraitz N. Zumeta[2], José J. Pizarro[2,3], Itziar Alonso-Arbiol[4]

[1] Department of Basic Psychological Processes and Development, University of the Basque Country, [2] Department of Social Psychology, University of the Basque Country, [3] Escuela de Psicología, Universidad Católica del Norte, [4] Department of Clinical and Health Psychology and Research Methods, University of the Basque Country

ABSTRACT: The neo-Durkheimnian model suggests that feedback and emotional communion between participants during a collective gathering (i.e., perceived emotional synchrony: PES) is one of the key mechanisms of collective processes. This shared emotional experience gives rise, in turn, to more intense emotions, this being one of the explanatory models of the positive psychological effects of collective participation. Through a quasi-longitudinal design of three measurement-times ($N = 273$, 65.9% women; age: 18-70, $M = 39.43$, $SD = 11.64$), the most massive social mobilization that is celebrated in favor of the Basque language in the Basque Country (*Korrika*) was analyzed. Repeated measures and sequential mediation analyze supported the model. The effect of participation on social integration was mediated by the increase in emotions of enjoyment through PES; the effect on social acceptance, social contribution, and social actualization was mediated by increased kama muta through PES; the effect on collective empowerment was mediated by the increase in self-transcendent emotions through PES; and the effect on remembered well-being was partially mediated by PES. Finally, it was also verified for the first time that the effect of participation on social integration, social acceptance and social actualization was maintained through PES (but not through emotions) for at least six to seven weeks after the event ended. Also, it is concluded that Kama muta is a relevant emotion during collective gatherings.

1. Introduction

The social isolation measures taken in different countries around the world to deal with the pandemic have been associated with higher rates of stress, anxiety and depression (Bueno-Notivol *et al.*, 2021; Marroquín *et al.*, 2020). Fortunately, these measures have been lifted in most countries of the world, allowing the population to return to social interactions or collective gatherings —social situations in which two or more people meet in one place with a common goal (McPhail & Wohlstein, 1983)—which

* **Corresponding author / Harremanetan jartzeko:** Itziar Alonso-Arbiol. Faculty of Science and Technology. University of the Basque Country (UPV/EHU). PO Box 644 (48080 Bilbao, Spain). e-mail: alonsoarbiol@gmail.com. **Contributed by** Itziar Alonso-Arbiol **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Zabala, *Jon et al.* «*Korrika*, running in collective effervescence through the Basque Country: a model of collective processes and their positive psychological effects» *Front. Psychol.* 14:1095763 (https://doi.org/10.3389/fpsyg.2023.1095763)

have been shown to be important for people's psychological health (Dimmock *et al.*, 2021; MacDonald & Hülür, 2021). Although it has been found that collective participation has a wide variety of positive psychological effects, more efforts have to be made to explain how these effects are generated during collective gatherings.

Through a quasi-longitudinal design, this work aims to analyze the positive psychological effects of participation in *Korrika*, the most important and massive event held in the Basque Country in favor of *euskera* (Basque language). Based on the neo-Durkheimnian model of collective processes proposed by Páez *et al.* (2015) we will analyze the mechanisms involved in the development of these effects, as well as their durability over time.

1.1. Collective Gatherings and Social Movements

Massive collective gatherings contain a high emotional and symbolic charge, and can profoundly mark the individual and collective life of people (Collins, 2004; Durkheim, 1912/1995). A series of studies, in line with Durkheim's theory (1912/1995), have shown that participation in collective gatherings has a wide variety of positive psychological effects. For example, it has been found that participation in collective gatherings is related to an increase in social cohesion —e.g., integration and social identity, perceived social support and solidarity—, leading to positive effects on social beliefs, such as the benevolence of people and society, and to an increase in empowerment at the individual level —e.g., self-esteem and life satisfaction— and, at the collective level, to a higher collective self-esteem and perceived collective efficacy (Bouchat *et al.*, 2020; Drury & Reicher, 2009; Khan *et al.*, 2016; Páez *et al.*, 2011, 2013, 2015; Tewari *et al.*, 2012; Wlodarczyk *et al.*, 2017a; Zumeta *et al.*, 2016a).

Among these psychological effects, collective identity has been highlighted as one of the most important factors that predict collective action, which in turn predicts collective efficacy, another of the most important factors of social mobilizations (Agostini & van Zomeren, 2021; van Zomeren *et al.*, 2008). Therefore, collective gatherings could be considered important fuels for social mobilizations (Pizarro *et al.*, 2022).

Finally, collective gatherings have also been shown to be important resources to counteract the negative effects of painful and traumatic events such as attacks, natural disasters, or social isolation during pandemic. In the face of these events, the strengthening of solidarity, support, and social cohesion during collective gatherings helps in the post-traumatic process (Páez *et al.*, 2007; Pelletier, 2018; Wlodarczyk *et al.*, 2016, 2017b, Zlobina & Dávila, 2022).

However, it is still necessary to clarify the psychological mechanisms involved in producing these effects. Because mere participation in collective gatherings does not necessarily provoke particular effects (see Thonhauser, 2022; see also Collins, 2004), outcomes generated in these instances of social life should involve the activation of one or more mechanisms.

1.2. Collective Effervescence

Durkheim's theoretical model (1912/1995) on collective processes provides a series of key elements to answer this question. Durkheim, who was interested in knowing what makes people stick together in society, developed the theory of collective effervescence. He maintained that collective gatherings are fundamental pillars for the individual and collective life of people. These gatherings, periodically held, fulfill the role of recreating the social group and reviving shared values. Participation in such gatherings fosters social cohesion and leaves participants with a renewed confidence in society and a sense of energy on an individual and collective level.

For these positive psychological effects to take place, collective effervescence is necessary. Durkheim described collective effervescence as a collective emotional exaltation that arises from contagion and emotional feedback between the people gathered. Each emotion resonates and feeds back among the participants, reaching a point of exaltation and emotional communion. Regardless of the type of af-

fect that predominates during the collective gathering —for example, sadness at a funeral, joy at a celebration or anger at a demonstration—, the positive psychological effects of collective participation are the result of emotional communion, that is, from the shared emotional experience and the feelings of unity derived from it (Páez *et al.*, 2015).

1.3. Perceived Emotional Synchrony

Páez *et al.* (2015), intending to empirically contrast Durkheim's approaches, proposed the Perceived Emotional Synchrony (PES) as a measure of collective effervescence; it is described "as an emotional experience felt by participants during group gatherings, involving a sense of togetherness" (Wlodarczyk *et al.*, 2023, p. 4756). Various meta-analyses have shown that experimentally induced synchronous behaviors foster perceived social bonding and positive affect (e.g., Mogan *et al.*, 2017; Rennung & Göritz, 2016). Synchronous behaviors such as singing, dancing, and repeating the same gestures and shouts are common during gatherings. These synchronous and symbolic behaviors during collective meetings awaken an emotional energy that intensifies and feeds back among the participants, giving rise to a shared emotional experience, reciprocal empathy and collective consciousness, that is, collective effervescence or PES (Collins, 2004; Durkheim, 1912/1995; Páez *et al.*, 2015; Wlodarczyk *et al.*, 2020, 2023).

Through this neo-Durkheimian model proposed by Páez *et al.* (2015) and subsequent studies that have followed this line of research, it has been found that the psychological effects of collective gatherings do not take place by mere collective participation, but rather that these effects are facilitated by the PES (Bouchat *et al.*, 2020; Castro-Abril *et al.*, 2021; Páez *et al.*, 2015; Pelletier, 2018; Pizarro *et al.*, 2019; Pizarro *et al.*, 2022; Wlodarczyk *et al.*, 2020, 2023; Zumeta *et al.*, 2016b, 2020). These studies suggest that the PES is one of the most important mechanisms of collective processes. The emotional communion generated from the emotional feedback between the participants gives rise, in turn, to more intense emotions during the collective gathering. As a result of this process, the participants return to their daily lives revitalized, that is, with a greater sense of social cohesion or integration, trust and empowerment at the individual and collective level (Páez *et al.*, 2015; Pizarro *et al.*, 2022; Wlodarczyk *et al.*, 2020, 2023; Zumeta *et al.*, 2020, 2021).

In the quasi-longitudinal study by Wlodarczyk *et al.* (2023), the relationships between PES, enjoyment emotions and self-transcendent emotions and their effects in a folk ritual were analyzed through structural equation models. PES acted as a direct predictor of increased self-transcendent emotions and emotions of enjoyment, as well as increased social integration, beliefs of a benevolent world and collective self-esteem. However, indirect effects of PES were also found through self-transcendent emotions on benevolent world beliefs and ingroup solidarity, as well as through enjoyment emotions on social integration.

These results show that the self-transcendent emotions and the emotions of enjoyment fostered by the PES during a gathering can also facilitate the effects of collective participation (i.e., the effects that the participants carry with them after leaving the collective situation). Based on this process, and on the existing literature on the effects of collective effervescence, we may refer to *proximal and distal effects* of PES (Pizarro *et al.*, 2022; Wlodarczyck, 2020; Zumeta *et al.*, 2020).

1.3.1. Positive Emotions and Kama Muta During Collective Effervescence

The most immediate or proximal effects of PES are those that occur during the same effervescence or collective situation, for example, positive self-transcendent emotions (Pizarro *et al.*, 2022). Self-transcendent emotions (e.g., hope, inspiration, and feeling grateful and amazed) bring a person out of their reverie by making them more receptive to stimuli from the social and natural world around them (Fredrickson, 2009; Haidt, 2006; Van Cappellen & Rimé, 2014). These emotional states induce an experience of self-transcendence, that is, a mental state in which the salience of the self decreases or in which feelings of connection with

other people or entities increase (Yaden *et al.*, 2017). This transcendence or connection induced by self-transcendent emotions facilitates, for example, the acceptance of people and the society they make up, as well as interest in their well-being (Aquino *et al.*, 2011; Pizarro *et al.*, 2021; Stellar *et al.*, 2017). Likewise, specifically in the case of the feeling of being moved and hope, these are significant factors related to collective efficacy and the action implied in social movements (Cohen-Chen & Van Zomeren, 2018; Landmann & Rohmann, 2020; Wlodarczyk *et al.*, 2017a; Zumeta *et al.*, 2021).

Another possible proximal effect of the PES, and to which a self-transcendental character can be attributed, is *kama muta*. Kama muta is an emotion used to refer to being moved by love of others (Fiske *et al.*, 2017; Zickfeld *et al.*, 2019). It is based on the horizontal sharing of relationships and arises from the sudden activation of these relationships in which people perceive each other as an equal in some essential aspect that they share (Fiske, 1992; Fiske *et al.*, 2017; Seibt *et al.*, 2019). For example, reunions, reconciliations and acts of friendship are some of the experiences that can evoke kama muta (Alfaro-Beracoechea & Contreras-Tinoco, 2021). During the emotional experience of kama muta, people feel mutual love, identification, solidarity, pity, kindness and devotion, promoting the desire to commit more strongly in those relationships (Fiske *et al.*, 2017; Zickfeld *et al.*, 2019).

We currently have no evidence of previous studies that have analyzed the role of kama muta during collective gatherings; however, it has been shown that one of the evocative experiences of kama muta are the big collective gatherings (Alfaro-Beracoechea & Contreras-Tinoco, 2021). In addition, some works suggest that social movements can arouse kama muta and induce a sense of moral commitment towards other people, for example, arousing the motivation to support causes related to a group (Fiske *et al.*, 2017; Landmann & Rohmann, 2020; Lizarazo *et al.*, 2022; Seibt *et al.*, 2019). We believe that PES can be a strong trigger for kama muta, especially during a collective gathering in which people come together for a common cause, as in the case of this study.

Lastly, the emotions of enjoyment, such as joy and fun, are expressed collectively during collective gatherings; although they cannot be considered self-transcendent emotions, in some studies they have also been related to feelings of unity and increased social integration (Novelli *et al.*, 2013; Wlodarczyk *et al.*, 2023).

1.3.2. Well-being and Collective Empowerment After Collective Effervescence

Distal effects are the effects that last or extend beyond the collective effervescence (i.e., the effects that the participants take with them into their daily lives after leaving the gathering). As mentioned above, people's evaluation of social cohesion, social beliefs and empowerment usually improve after collective gatherings (Pizarro *et al.*, 2022). These aspects are often closely related to people's psychological health, for example, to *social well-being*.

Social well-being is the evaluation that people make of the circumstances and functioning they have within society (Keyes, 1998), and is made up of five dimensions: social integration, social acceptance, social contribution, social actualization and social coherence. *Social integration* refers to the feeling of belonging and being accepted by other people who constitute their social reality and the degree of similarity perceived with them. Social belonging is a good indicator of people's psychological health, and a factor that protects against anxiety, stress and depression (Páez & Oyanedel, 2021; Postmes *et al.*, 2019; Zabala *et al.*, 2020). *Social acceptance* refers to the positive view of human nature and trust in others. According to Keyes (1998), this belief about people's benevolence is also an important indicator of people's health, since psychologically healthier people tend to attribute their own benevolence and personal acceptance —acceptance of bad things and good— also to other people. This trust placed in other people is accompanied by *social contribution*: the belief that oneself is also a useful member of society. This dimension is based on perceived self-efficacy and sense of control (Bandura, 1997; Keyes, 1998), relevant factors in people's psychological health. As for *social actualization*, this

refers to the belief that society progresses in a beneficial direction for people. Lastly, the psychologically healthier people, in addition to placing trust in oneself, in the people and the society they make up, also tend to worry about knowing and understanding the social life that surrounds them; this is what Keyes (1998) named *social coherence*, the feeling that one is able to understand what is happening around him.

After collective gatherings, people's evaluation of quality of life tends to improve even in such personal aspects as self-esteem, sense of control and autonomy, as well as meaning and satisfaction with life (Páez *et al.*, 2015; Wlodarczyk *et al.*, 2023; Zumeta *et al.*, 2016a), aspects included in the *remembered well-being* of Hervás & Vázquez (2013).

Finally, collective gatherings in addition to being beneficial at the individual level, can also be beneficial at the collective level, and increase perceived collective efficacy —the shared belief of a group in its ability to organize and execute the actions required to reach certain levels of achievement (Bandura, 1997)— (Zumeta *et al.*, 2016a, 2020, 2021). When it comes to minority or discriminated groups, and this perceived efficacy is aimed at counteracting existing power relations and influencing social change, it is referred to as *collective psychological empowerment* (Drury & Reicher, 2009), which may also be strongly related to the well-being of people (Zabala *et al.*, 2020).

1.4. *The Durability of the Effect of Participation*

Durkheim stresses the importance of the regular practice of collective gatherings since the effect of participation is diluted over time. The durability of the effect of collective participation has been analyzed on very few occasions. In the study by Pizarro *et al.* (2019) the effect of participation in a mindful-dancing program dissipated within one week of participation. In contrast, in other studies, the effect of participation was maintained for at least three weeks (Páez *et al.*, 2007, 2015; Rimé *et al.*, 2010) and four weeks after participation (Khan *et al.*, 2016; Páez *et al.*, 2011; Tewary *et al.*, 2012). In the study by Bouchat *et al.* (2020) the distal effects of the PES during the most massive event of scouts were analyzed. PES during the encounter predicted scores for some of the distal effects after 10 weeks of participation. In the present study, we will analyze the effects of participation up to six-seven weeks after having participated.

1.5. *Current Study: Korrika, the Claim of the Language and the Culture on the Territory*

Korrika[1] is a crowded race of a leisure nature that every two years crosses the Basque Country (a region currently split between Spain and France) for 11 days without interruption. The objective of this ritual is to raise funds to promote Euskera, a minority[2] and isolated language spoken in the Basque Country, whose origin is a mystery and that continues to challenge many scholars to decipher it. Basque is an essential element, almost sacralized for the identity of Basque speakers, and is estimated that around six hundred thousand people participate in Korrika, making this ritual a highly emotional and indescribable celebration for many participants (del Valle, 1998).

Korrika contains all the elements of a collective ritual and therefore the conditions for the collective effervescence to arise (i.e., PES), as well as the emotions of enjoyment, self-transcendent emotions and kama muta. The ritual, as a whole, is made up of symbolic behaviors that represent and claim the territory of the Basque language and the entire Basque culture. Each group of participants travels a section of the route while carrying a baton with a secret message inside that is revealed in the massive celebration at the end of the route. This act

[1] Korrika is organized by the Literacy and Euskaldunization Coordinator (AEK, in its acronym in Basque) and also seeks to raise funds for its network of Basque teaching centers (euskaltegiak). AEK is the most important organization in the field of euskaldunization and literacy of adults, and the only one that develops its activity throughout Basque Country, above administrative divisions. https://www.aek.eus/es-es/servicios/korrika#

[2] Basque speakers whose mother tongue is Euskera account for 15.35% of the Basque Country's population However, 29.62% of the inhabitants of the Basque Country have acquired language ability in Euskera, even though it is only practiced by 12.6% in public life (Sociolinguistic Cluster, 2020).

of going through the entire territory with a baton that passes from hand to hand symbolizes the legacy of Basque —that Basque does not stop, does not disappear as long as there are runners (i.e., speakers) who practice it—. Both at the beginning and at the end, as well as throughout the race, repetitive, stereotyped and synchronous behaviors are expressed, such as clapping, singing and shouting. This is accompanied by music and symbolic elements that express shared values, such as flags and traditional elements in clothing (del Valle, 1998). This great social mobilization around the Basque language is likely to be explained by the strong identity and collective empowerment of the Basque society. However, we believe that between these factors and the collective effervescence there is mutual feedback that strengthens the social movement, and that is also crucial.

1.6. Objectives and Hypotheses

The present study has three objectives:

1. To test whether participation in the Korrika collective ritual fosters social well-being, remembered well-being and collective empowerment (dependent variables).
2. To examine if the effect (if any) of the quality of participation on the dependent variables is mediated by the PES and its proximal effects (emotions of enjoyment, self-transcendent emotions and kama muta) in this sequence, and after controlling the effect of the dependent variables before having participated in Korrika.
3. To analyze the durability of the effect of participation up to six-seven weeks after having participated, and analyze the effect of the same mediator's variables (PES and proximal effects) on those dependent variables where the effect of the participation still endures.

Related to these objectives, the following hypotheses were formulated:

— H1. Participants are expected to express higher social well-being, remembered well-being and collective empowerment after having participated in Korrika.
— H2. The effect of the quality of participation on the dependent variables is expected to be mediated by the PES, or by the increase in the emotions of enjoyment, self-transcendent emotions and kama muta through the PES.
— H3. The durability of the effect of participation up to six-seven weeks after Korrika has finished will be explored. No assumptions are made in this regard. However, if the effect of participation persists on some dependent variables, it is expected that the PES or its proximal effects will predict positively that effect.

2. Materials and Methods

2.1. Participants

Three questionnaires were administered in three times. The number of people who answered the questionnaire of Time 1 was 748 (63.1% women and 36.9% men), aged between 18 and 73 years (M = 39.28; SD = 12.13). From that sample only 404 people (66.1% women and 32.2% men), aged between 18 and 73 years (M = 39.79, SD = 11.74) answered also the questionnaire of Time 2. Finally, from that sample (T1-T2), 273 people answered the questionnaire of Time 3 (65.9% women and 33.01% men), aged between 18 and 70 years (M = 39.43, SD = 11.64). All the participants were citizens from the historical territory of the Basque Country.

2.2. Procedure

Given that our objective in this study was to test hypotheses of causal relationships, and despite the limitations posed by quasi-experimental studies, and the reality that we intended to study, we believed that it was the most appropriate design to achieve our objectives. Data was collected in three times. The first survey was administered three weeks before starting Korrika (Time 1: T1). The second survey was administered between the first and seventh days after participating in Korrika (Time 2: T2). Finally, the third survey was administered between six and seven weeks after Korrika had finished (Time 3: T3). The data of the dependent vari-

ables —i.e., measures of the distal effects— (social well-being, remembered well-being and collective empowerment) were collected at T1, T2 and T3. Data on quality of participation, PES, enjoyment emotions, self-transcendent emotions, and kama muta were collected at T2.

Once the approval of the Ethics Committee of the university to which the authors belong was obtained (M10_2019_004), there were collected the emails of those people who intended to participate in Korrika 2019 and in the research. *Alfabetatze eta Euskalduntze Koordinakundea* (AEK; the association that organizes the Korrika event) collaborated with the research team by sending an email invitation to all Korrika collaborators, requesting their voluntary participation in the research. On the other hand, researchers went to the capitals of the Basque Country and collected the emails from those who had been willing to participate in Korrika and in this research. Participants answered all questionnaires online through the *Qualtrix XM* platform.

2.3. *Instruments*

Quality of participation (Páez *et al.*, 2015). The involvement of the participants in Korrika was measured with three items on a 7-point scale (1 = *Not at all*, 7 = *Very much*). For example: "How intense was your participation? ". Obtained reliability indexes were satisfactory (α = .83; Ω = .84).

Perceived Emotional Synchrony (PES; Páez *et al.*, 2015, brief version by Wlodarczyk *et al.*, 2020). Through four items, the extent to which the participants experienced collective effervescence was analyzed on a 7-point scale (1 = *Not at all*, 7 = *A lot*). For example: "We felt a strong shared emotion". Obtained reliability indexes were satisfactory (α = .92; Ω = .92).

Enjoyment and Self-Transcendent Emotions (Modified Differential Emotions Scale-MDES; Fredrickson, 2009). The extent to which participants felt two enjoyable emotions (amusement/entertainment and joy/happiness) and four self-transcendent positive emotions (wondered, grateful, inspired, and hopeful) was analyzed on a 5-point ordinal scale (1 = *Not at all*, 2 = *A little bit*, 3 = *Moderately*, 4 = *Quite a bite* y 5 = *Extremely*). For example: "What is the most joyful, glad, or happy you felt?" or "What is the most grateful, appreciative, or thankful you felt?". Obtained reliability indexes were satisfactory: *r* = .53 (α = .68) and α = .78 (Ω = .78) respectively.

Kama muta (KAMUS; Zickfeld *et al.*, 2019). The extent to which participants had experienced kama muta was analyzed on a 7-point scale (1 = *Not at all*, 7 = *Very much*). The data of the evaluation dimensions, of four items, were analyzed; for example: "An exceptional sense of closeness appear" and of motivation (also made up of four items); for example: "I wanted to hug someone". Obtained reliability indexes were satisfactory, α = .92 (Ω = .93).

Social Well-Being (SWB; Keyes, 1998, adapted by Blanco & Díaz, 2005). Using a 5-point scale (1 = *Strongly disagree*, 5 = *Strongly agree*), the following dimensions were analyzed: social integration (e.g., "I feel close to other people in my community"), social acceptance (e.g., "I think that people are basically good"), social contribution (e.g., "I have something valuable to give the world") and social actualization (e.g., "Society is making progress, getting better"). Each dimension was made up of three items, as in the original version of 15 items (Keyes, 2002). We did not analyze the social coherence dimension. The reliability indices in social integration ranged from α = .69 (Ω = .70) at T1 and α = .84 (Ω = .85) at T3. In social contribution range was between α = .75 (Ω = .76) and α = .81 (Ω = .82), and in social actualization between α = .83 (Ω = .83) and α = .85 (Ω = .86). For social acceptance, reliability indices were lower: between α = .48 (Ω = .50) at T1 and α = .62 (Ω = .64) at T2.

Remembered Well-Being (The Pemperton Happiness Index-PHI; Hervás & Vázquez, 2013). Through 10 items, the participants' recalled well-being was analyzed on a 10-point scale (1 = *Strongly disagree*, 10 = *Strongly agree*). Data on general well-being (e.g., "I am very satisfied with my life"), eudaimonic (e.g., "I think my life is useful and worthwhile") and hedonic (e.g., "I enjoy a lot of little things every day") was collected. However, the hedonic dimen-

sion was removed from the analysis, because the only inverse item (i.e., "I have a lot of bad moments in my daily life") showed a considerably small factor loading, and this negatively affected the fit indices. Indeed, we found that a lot of participants responded to this item in the opposite direction in which the answer should have been given. Both dimensions were treated as an only measure of well-being. Obtained reliability indexes were satisfactory, ranging between α = .90 (Ω = .90) in T1 and α = .93 (Ω = .93) in T3.

Collective Empowerment (Collective Efficacy Questionnaire for Sports-CEQS; Martínez *et al.*, 2011). The group efficacy (that is, as Basque speakers) perceived by the participants was measured with four items on a 10-point scale (1 = *Not at all capable*, 10 = *Very capable*). For example, "To carry out actions". In addition, we added two items elaborated *ad hoc* to capture Drury and Reicher's (2009) concept of collective psychological empowerment, specifically: "To promote social change" and "To achieve common goals". Obtained reliability indexes were satisfactory, ranging between α = .89 (Ω = .89) in T1 and α = .91 (Ω = .91) in T3.

2.4. *Data analysis*

We conducted ANOVA repeated measures analysis in order to test Objectives 1 and 3. To analyze the effects of the mechanisms involved in changes from T1 to T2 (i.e., Objective 2), we conducted several structural equation models using the quality of participation as the main predictor, and a sequential mediation approach from PES to the emotional aspects (i.e., emotions of enjoyment, self-transcendent emotions and kama muta). Finally, for each dependent variable we included the score in T1 to calculate specifically possible changes due to having participated in the ritual. Finally, to analyze the durability of the effect of the mechanisms until T3 (i.e., Objective 3), we first performed linear regressions on the dependent variables where the effect of participation persisted, to then carry out a series of structural equation models that were most consistent with the previous results and the fit indicators. For these models, no outliers were excluded, and the proposed theoretical model was compared with other alternative models, which was shown to be more adequate both theoretically and statistically (See Supplementary Table S2 and S3).

To evaluate the fit of the models, we included the chi-square test[3] together with CFI (Comparative Fit Index) and TLI (Tucker-Lewis Index), whose values over 0.90 are considered acceptable. Likewise, RMSEA (Root Mean Square Error of Approximation) and SRMR (Standardized Root Mean Square Residual) were taken into account, whose values below .05 indicate a good model, and with values between .05 and .08 indicating that the model is reasonably good (Hu & Bentler, 1999). The applied estimation procedure was *Maximum Likelihood* and all analyses were conducted through JASP 0.15.

Concerning the participants, Objective 1 and Objective 2 were accomplished with the sample of participants who responded to T1 and T2. Instead, Objective 3 was accomplished with the sample of participants who had answered the questionnaires in T1, T2 and T3. There were no statistically significant differences between these groups neither in gender, nor in age, nor in the dependent variables before having participated in Korrika (See Supplementary Table S1).

3. Results

3.1. *Descriptive and preliminary analyses*

Means and standard deviations, correlations, and collinearity statistics are shown in the online supplemental materials (Supplementary Tables S2-S4). All the analyzed variables in T2 were related significantly and positively. Correlation ranged from $r = .16$ between the quality of participation and social actualization to $r = .70$ between enjoyment and self-transcendent emotions; all $ps < .001$ (See Supplementary Table S3)[4].

[3] The first number between parentheses refers to the sample and the second number refers to the number of estimates (See Figures 1-6).

3.2. *Effect of Participation on Dependent Variables*

The results of the repeated measures ANOVAs (see Table 1) show that the scores of the dependent variables in T1 increase significantly in T2, except in the case of remembered well-being. Furthermore, it is worth mentioning that the other effect sizes —except for collective empowerment— are satisfactory, especially in the case of social actualization. In general, we observed an increase in the scores of the dependent variables from T1 to T2; therefore, Hypothesis 1 was confirmed.

Table 1
Results of repeated measures of dependent variables in T1 and T2

Dependent Variables	T1 *M* (*SD*)	T2 *M* (*SD*)	*F*(1, 403)	*p*	η_p^2
Social integration	3.98 (0.60)	4.15 (0.64)	31.626	.001	.07
Social acceptance	3.20 (0.60)	3.41 (0.67)	37.390	.001	.09
Social contribution	3.60 (0.67)	3.74 (0.73)	19.377	.001	.05
Social actualization	2.52 (0.77)	2.94 (0.83)	134.870	.001	.25
Remembered well-being	7.89 (1.10)	7.97 (1.17)	3.303	.070	.01
Collective empowerment	7.78 (1.27)	8.02 (1.34)	14.143	.001	.03

3.3. *Sequential Mediation Analysis*

We tested whether the effect of quality of participation (QP) on the dependent variables was mediated by increased enjoyment emotions, self-transcendent emotions, and kama muta through PES. The structural equation models can be seen in Figures 1-6 (the coefficients shown in the figures are standardized coefficients and the solid arrows indicate that the coefficient is statistically significant). To check the sequential mediation, we calculated the indirect effects of the QP through PES and the proximal effects (emotions of enjoyment, self-transcendent emotions and kama muta) in this sequence (standardized effects and confidence intervals are presented).

3.3.1. Emotions of Enjoyment, Self-transcendent Emotions and Kama muta

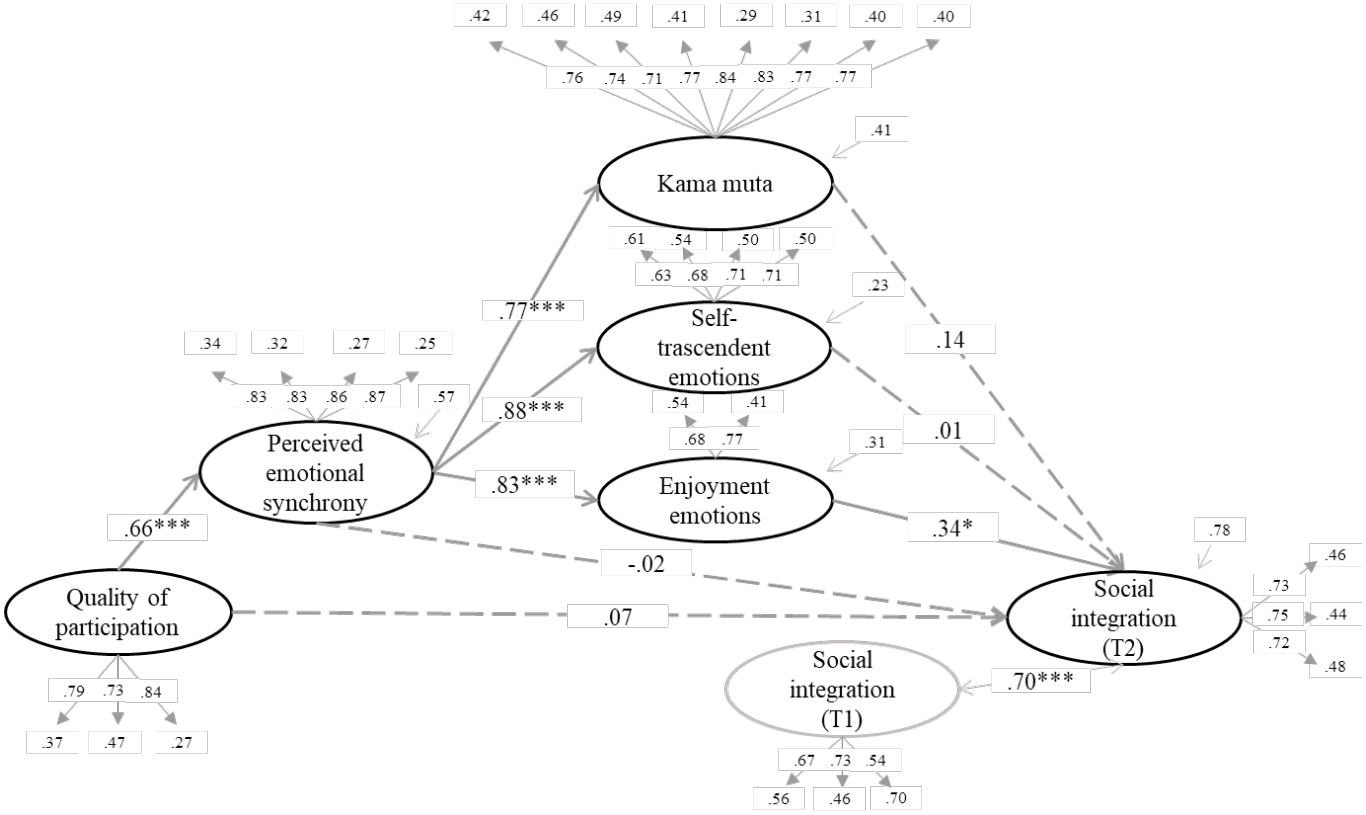
As it can be seen in all the figures mentioned above, the QP predicts the increase in PES and, in turn, PES acts as a direct predictor of the increase in emotions of enjoyment, of emotions of self-transcendent and kama muta. It is worth mentioning that in another alternative model enjoyment emotions did not predict PES, and that the betas of self-transcendent and kama muta emotions on PES were lower than the betas of PES on them (see Supplementary Table S3). This result further reinforces the proposed model.

Effects on Social Well-Being: Social Integration, Social Acceptance, Social Contribution and Social Actualization

In Figure 1 we can see that there is a direct effect of enjoyment emotions on social integration ($B = 0.34$, $SE = 0.14$, $p = 0.017$, 95% CI [0.060, 0.611]). In the ab-

[4] Collinearity analysis showed adequate indices (tolerance > 0.2, variance inflation factor < 4; Rovai *et al.*, 2013). See supplementary Table S5.

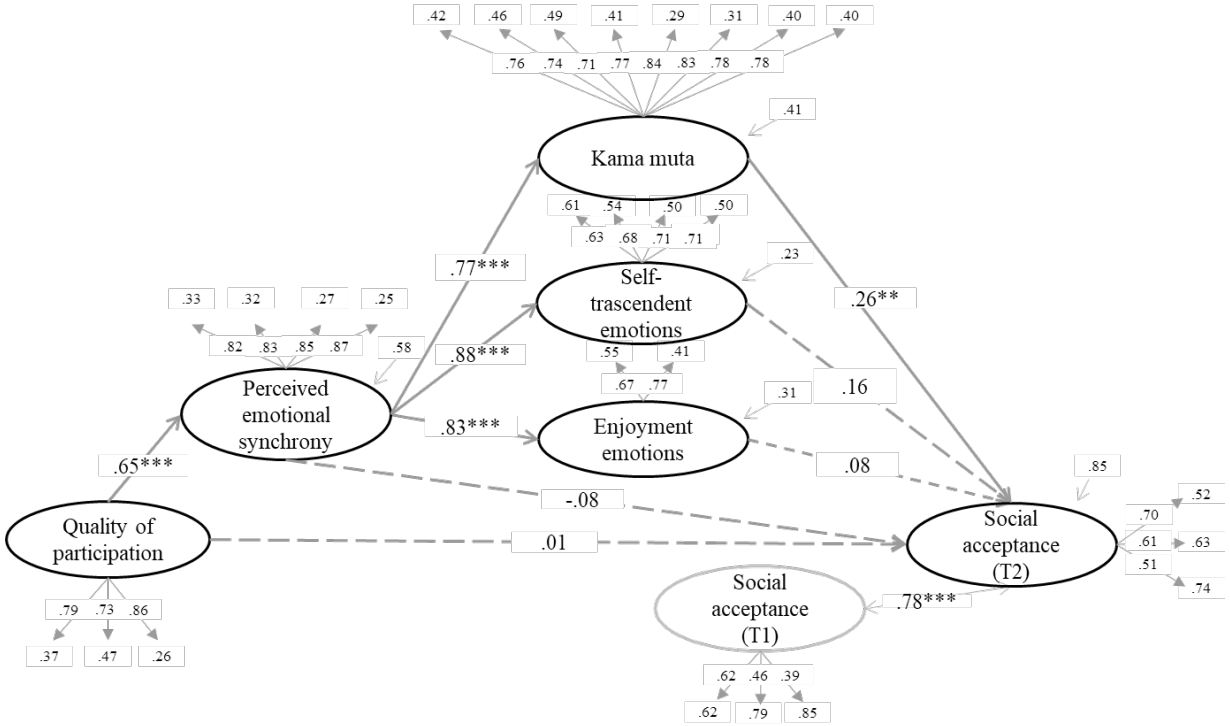
sence of the direct effect of PES on social integration, we analyzed the indirect effect of QP through PES and enjoyment emotions in this sequence. The indirect effect was statistically significant ($B = 0.19$, $SE = 0.04$, $p = 0.020$, 95% CI [0.015, 0.177]); therefore, we could state that the PES mediated the effect of the QP on social integration through of increased emotions of enjoyment.



Note. Model fit: $\chi 2 = (404, 313) = 811.975$, $p < 0.001$, CFI = 0.920, TLI = 0.910, RMSEA = 0.063, SRMR = 0.058.

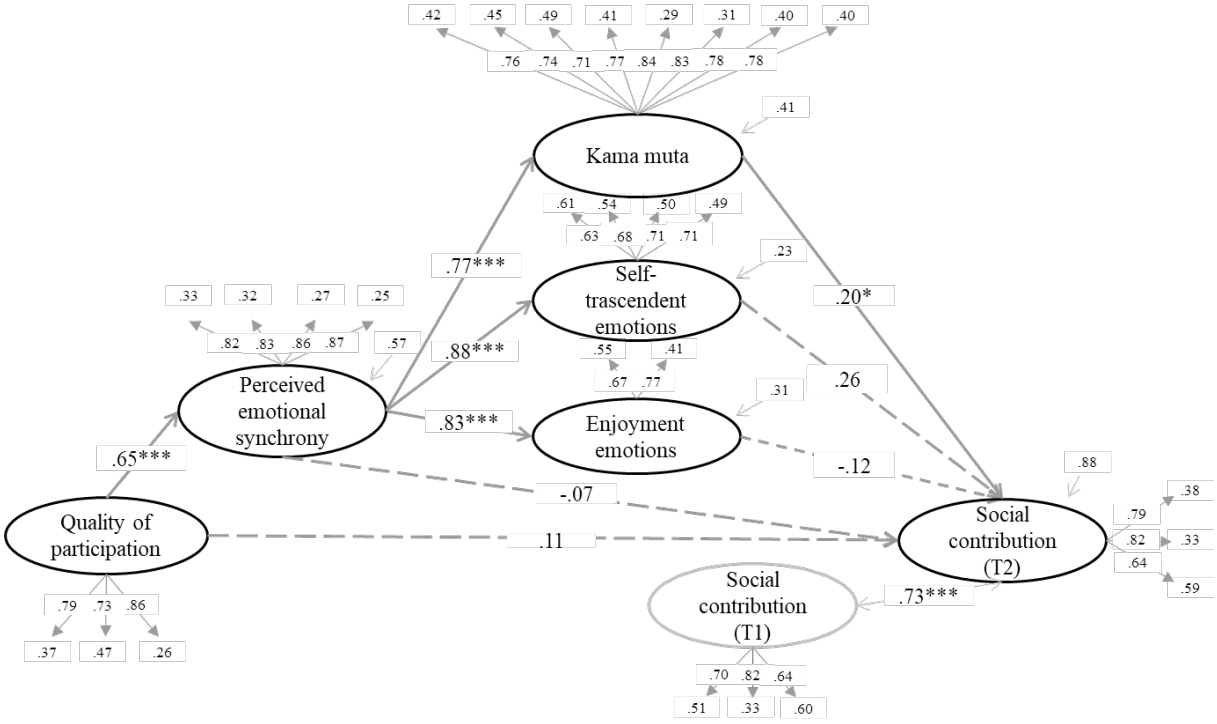
Figure 1. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on social integration controlling for pre-participation scores

In Figure 2, Figure 3 and Figure 4 we can see that there is a direct effect of kama muta on social acceptance ($B = 0.26$, $SE = 0.04$, $p = 0.008$, 95% CI [0.029, 0.190]), social contribution ($B = 0.20$, $SE = 0.04$, $p = 0.011$, 95% CI [0.022, 0.165]) and social actualization ($B = 0.27$, $SE = 0.04$, $p = 0.001$, 95% CI [0.058, 0.202]). Again, in the absence of the direct effect of PES, we calculated the indirect effect of QP through PES and kama muta, which was statistically significant in all three cases: on social acceptance ($B = 0.13$, $SE = 0.03$, $p = 0.009$, 95% CI [0.019, 0.137]), social contribution ($B = 0.10$, $SE = 0.03$, $p = 0.012$, 95% CI [0.014, 0.119), and social actualization ($B = 0.13$, $SE = 0.03$, $p = 0.001$, 95% CI [0.040, 0.147]). In these cases, PES mediated the effect of QP on the three variables through the increase in kama muta.



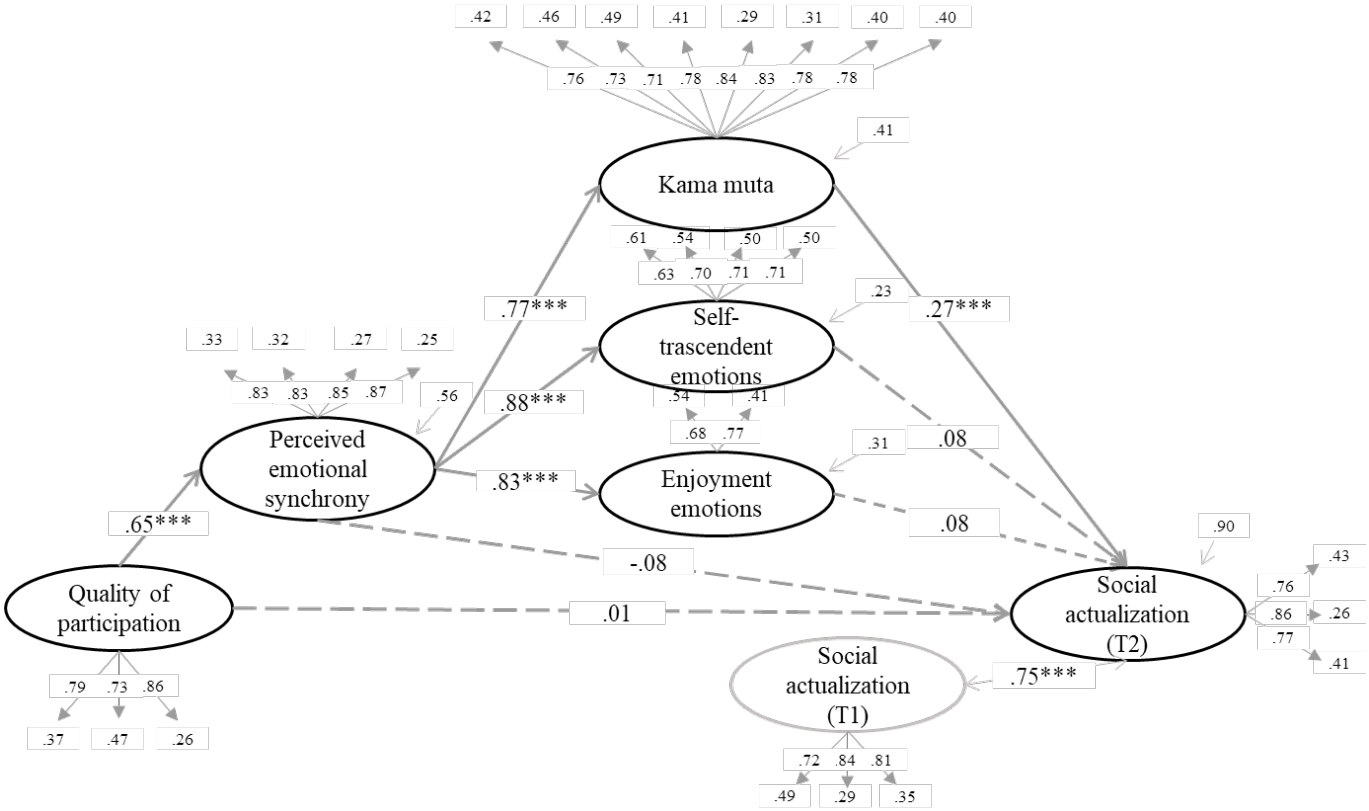
Note. Model fit: $\chi 2 = (404, 313) = 904.186$, $p < 0.001$, CFI = 0.900, TLI = 0.888, RMSEA = 0.068, SRMR = 0.060.

Figure 2. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on social acceptance controlling for pre-participation scores



Note. Model fit: $\chi 2 = (404, 313) = 881.122$, $p < 0.001$, CFI = 0.912, TLI = 0.901, RMSEA = 0.067. SRMR = 0.074.

Figure 3. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on social contribution controlling for pre-participation scores



Note. Model fit: $\chi 2 = (404, 313) = 816.599$, p <0.001, CFI = 0.925, TLI = 0.916, RMSEA = 0.063. SRMR = 0.056.

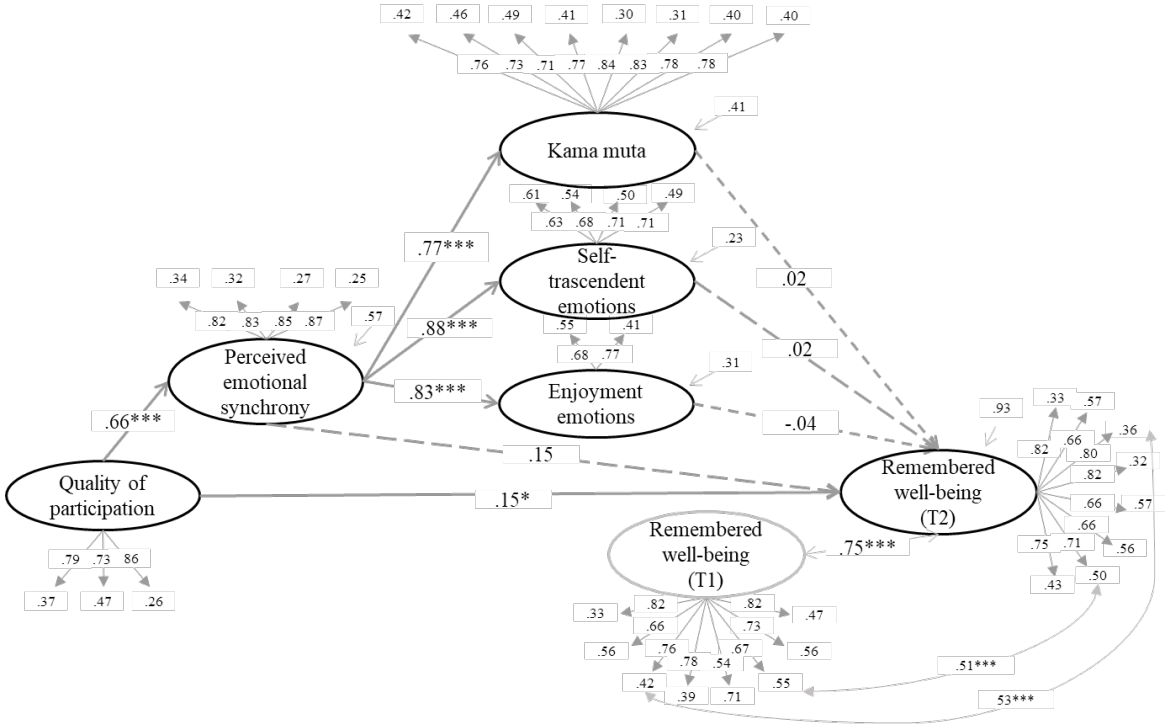
Figure 4. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on social actualization controlling for pre-participation scores

3.3.2. Effect on Remembered Well-Being

In Figure 5 we can see that there is a direct effect of the QP on remembered well-being (B = 0.15, SE = 0.08, p = 0.029, 95% CI [0.018, 0.333], and unlike the previous models, neither the PES neither did proximal effects mediate the effect of QP on remembered well-being. However, in the absence of direct effects of PES and proximal effects, we calculated its total indirect effect, which was statistically significant (B = 0.09, SE = 0.04, p = 0.010, 95% CI [0.026, 0.195]). Alternatively, we also performed a simple mediational analysis ($\chi 2$ = (404, 223) = 603.384, p < 0.001, CFI = 0.933, TLI = 0.923, RMSEA = 0.065, SRMR= 0.054), in which PES partially mediated the relationship between QP and remembered well-being (B = 0.13, SE = 0.06, p = 0.007, 95% CI [0.045, 0.286]). Therefore, we can say that PES and proximal effects also influenced the effect of QP on remembered well-being.

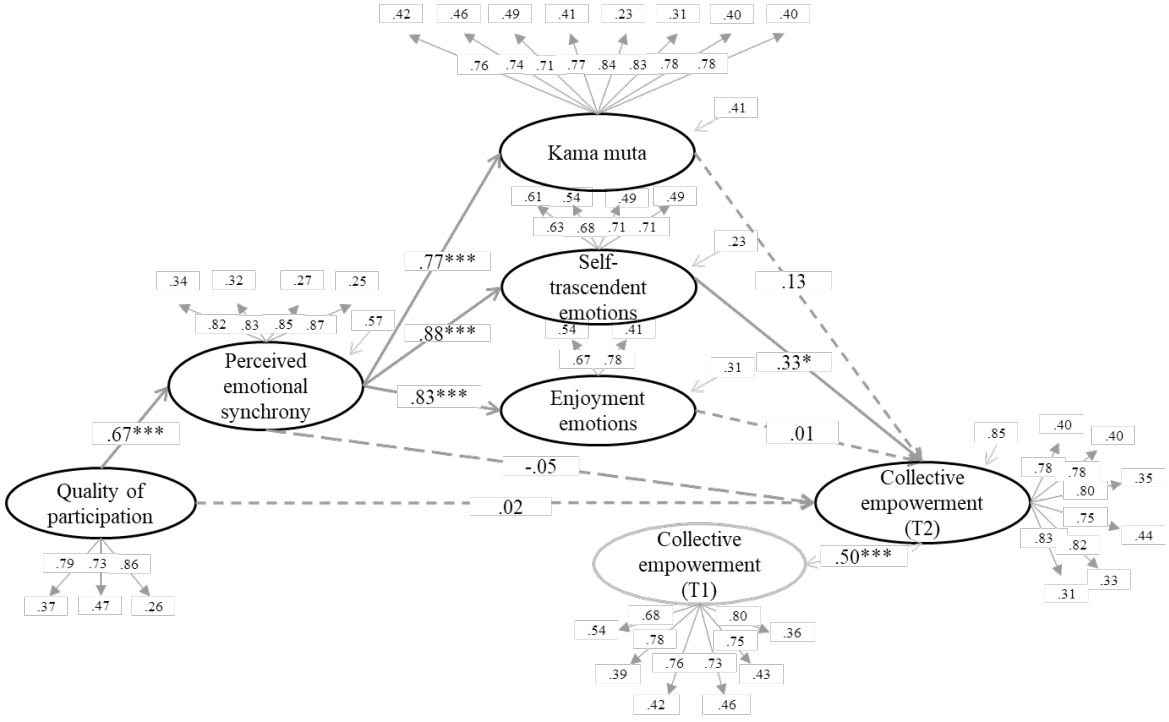
3.3.3. Effect on Collective Empowerment

In Figure 6 we can observe a direct effect of self-transcendent emotions on collective empowerment (B = 0.33, SE = 0.33, p = 0.031, 95% CI [0.065, 1.351]). In turn, we calculated the indirect effect of QP through PES and self-transcendent emotions, which also turned out to be statistically significant (B = 0.19, SE = 0.12, p = 0.033, 95% CI [0.020, 0.476]). In this case, we can say that PES mediated the effect of QP on collective empowerment through self-transcendent emotions.



Note. Model fit: $\chi 2 = (404, 616) = 1419.787$, $p < 0.001$, CFI = 0.914, TLI = 0.907, RMSEA = 0.057, SRMR = 0.056.

Figure 5. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on remembered well-being controlling for pre-participation scores



Note. Model fit: $\chi 2 = (404, 484) = 1237.669$, p <0.001, CFI = 0.912, TLI = 0.904, RMSEA = 0.062. SRMR = 0.062.

Figure 6. Perceived emotional synchrony and emotions as mediators of the effect of quality of participation on collective empowerment controlling for pre-participation scores

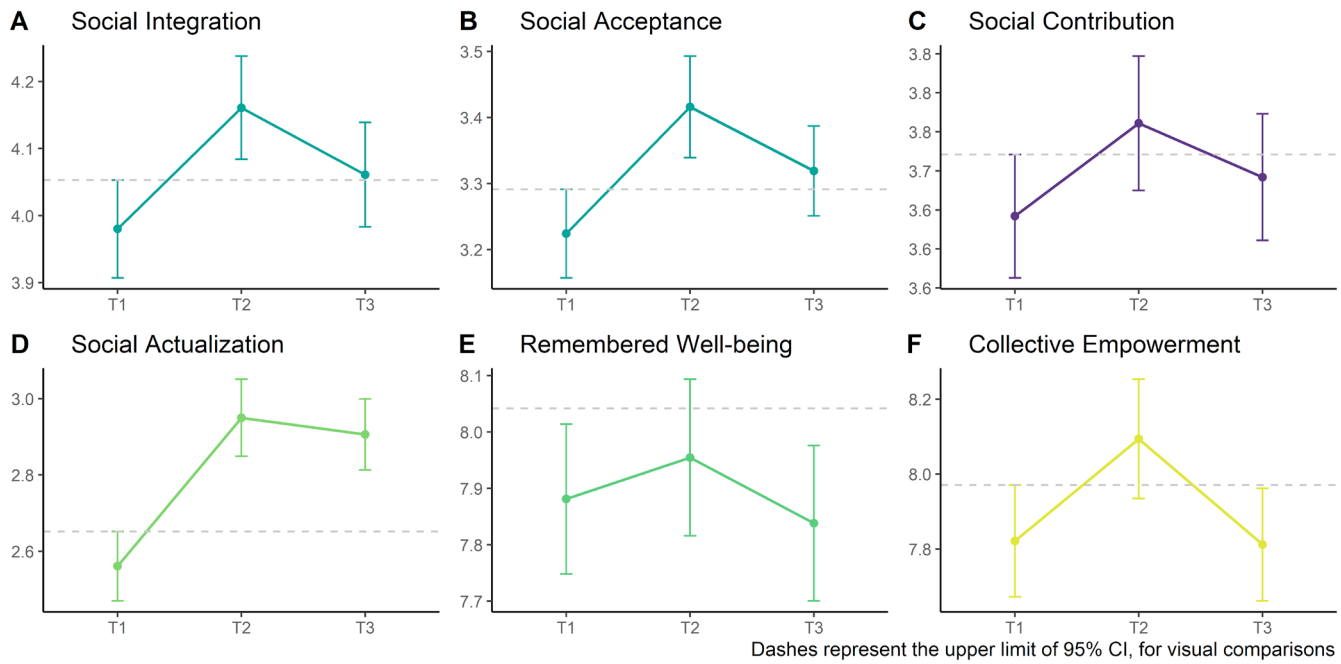


Figure 7. Results of repeated measures of dependent variables in T1, T2 and T3

3.4. *Durability of Effect of Participation*

As can be seen in Figure 7, scores of social contribution, remembered well-being, and collective empowerment return to their pre-participation status after six to seven weeks. However, scores for social integration, social acceptance and social actualization remain higher at T3 (i.e., 6-7 weeks after Korrika) than in T1.

Table 2

Multiple linear regressions of the effect of the quality of participation, of perceived emotional synchrony and of proximal effects on social integration, social acceptance and social actualization on T3, controlling for T1 scores

S. Integration^{T3}		S. Acceptance^{T3}		S. Actualization^{T3}	
	B		*B*		*B*
S. Integration^{T1}	.51***	S. Acceptance^{T1}	.43***	S. Actualization^{T1}	.58***
Quality of participation	−.08	Quality of participation	−.12	Quality of participation	−.06
PES	.13*	PES	.21**	PES	.16*
Enjoyment emotions	.03	Kama muta	.02	Kama muta	-.10
R^2	.31		.24		.35
F(4, 268)	29.357***		20.753***		36.689***

Note. *B*, standardized beta; *p < .05; **p < .01; ***p < .001.

In addition, as we can see in Table 2 the unique predictor variable of the scores of social integration, social acceptance and social actualization at T3 is PES, and not the proximal effects. These results were also supported by structural equation modeling (see Supplementary Figures 1-3), where the PES is the unique predictor of social integration[5] ($B = 0.28$, $SE = 0.06$, $p = 0.004$, 95% CI [0.059, 0.304]), social acceptance[6] ($B = 0.33$, $SE = 0.05$, $p = 0.010$, 95% CI [0.031, 0.219]), and social actualization[7] ($B = 0.22$, $SE = 0.07$, $p = 0.037$, 95% CI [0.009, 0.282]). Enjoyment emotions don't mediate the effect of PES on social integration ($B = -.02$, $SE = 0.12$, $p = 0.835$, 95% CI [−0.270, 0.218]), and kama muta neither mediates the effect of PES on social acceptance ($B = -.07$, $SE = 0.03$, $p = 0.543$, 95% CI [−0.080, 0.042]), nor on social actualization ($B = -.12$, $SE = 0.05$, $p = 0.214$, 95% CI [−0.148, 0.033]).

In summary, the changes in the dependent variables from T1 to T2 are positive and statistically significant except for remembered well-being. All QP effects were mediated by some proximal effect of PES, except for remembered well-being. However, even in this case, apart from being a total indirect effect of PES and proximal effects, in simple mediational analyses PES partially mediated the effect of QP; therefore, in these cases Hypothesis 2 was partially confirmed. Finally, the effect of PES — but not of the proximal effects— was maintained on social integration, social acceptance and social actualization until T3 (i.e., at least up to six-seven weeks after participation), also confirming Hypothesis 3.

4. Discussion

The Korrika collective ritual offers an excellent setting to investigate collective processes. This ritual symbolizes the vindication and legacy of the Basque language and culture over the historical territory and it contains a great emotional and symbolic charge, as well as all the essential elements for the collective effervescence to arise. The results of this semi-longitudinal and quasi-experimental study support the neo-Durkheimnian model of collective processes proposed by Páez *et al.* (2015).

On the one hand, it has been found that participation in Korrika effectively fosters social cohesion, renews trust in society and empowers individuals and groups. Specifically, we have found an increase in social integration after participation in Korrika, as well as positive effects on acceptance and social actualization. In turn, empowerment at the individual level, that is, social contribution, and empowerment at the collective level have also been favored. The trend increase in remembered well-being may be due to the fact that T1 scores were already high before participating in Korrika. In any case, previous studies show statistically significant effects on remembered well-being (Páez *et al.*, 2015; Zumeta *et al.*, 2016a; Wlodarczyk *et al.*, 2023).

On the other hand, in line with the neo-Durkheimian model, the results show that PES and its proximal effects (Páez *et al.*, 2015; Wlodarczyk *et al.*, 2020, 2023) facilitate these positive psychological effects. First, QP predicted PES, and PES was shown to be a predictor of enjoyment emotions, self-transcendent emotions, and kama muta. Thus, the proximal effects of PES facilitated all the effects of collective participation in Korrika, although partially in the case of remembered well-being.

In line with the work of Wlodarczyk *et al.* (2023), the effect on social integration was facilitated by the indirect effect of QP through PES and enjoyment emotions. Emotions such as fun and joy are experienced and expressed collectively during collective gatherings, and it seems that they can also lead to a greater connection with other people.

[5] Enjoyment emotions as mediator of the effect of PES on social integration in T3 controlling for pre-participation scores. Model fit: $\chi2 = (273, 49) = 111.382$, $p < 0.001$, CFI = 0.964, TLI = 0.952, RMSEA = 0.068, SRMR = 0.040.

[6] Kama muta as mediator of the effect of PES on social acceptance in T3 controlling for pre-participation scores. Model fit: $\chi2 = (273, 130) = 376.737$, $p < 0.001$, CFI = 0.910, TLI = 0.894, RMSEA = 0.083, SRMR = 0.060.

[7] Kama muta as mediator of the effect of PES on social actualization in T3 controlling for pre-participation scores. Model fit: $\chi2 = (273, 130) = 312.372$, $p < 0.001$, CFI = 0.945, TLI = 0.935, RMSEA = 0.072, SRMR = 0.041.

The effect on social acceptance, social contribution and social actualization was facilitated by the indirect effect of QP through PES and kama muta. Kama muta can be considered a self-transcendent emotion directly oriented to horizontal relationships, relationships in which people feel like neighbors or similar in certain essential aspects of the Self (Zickfeld *et al.*, 2019). In turn, it is likely that kama muta has also fueled the desire to take responsibility for the people and society they build (Fiske *et al.*, 2017; Landmann & Rohmann, 2020; Lizarazo Pereira *et al.*, 2022; Seibt *et al.*, 2019), probably strengthening social contribution, that is, the feeling of being valuable and capable of improving society. Undoubtedly, kama muta emerges as a relevant emotion during collective gatherings whose further study seems valuable.

Regarding the effect on collective empowerment, this was facilitated by the indirect effect of QP through PES and self-transcendent emotions —wondered, grateful, inspired and hopeful—. Hope and inspiration, as well as awe or amazement at the vastness of social movements (in this case Basque) can be powerful emotions in fostering a sense of being able to achieve common goals and change the collective situation (Cohen-Chen & Van Zomeren, 2018; Landmann & Rohmann, 2020; Wlodarczyk *et al.*, 2017a; Zumeta *et al.*, 2021). These data support the idea that collective effervescence can be an important fuel for social movements, since one of the most important factors that explain collective action is collective efficacy.

The effect on recalled well-being was mainly facilitated by QP and not by PES or its proximal effects. However, the indirect total effect of PES and proximal effects was statistically significant, suggesting that the effect on remembered well-being is not only explained by mere collective participation, but also by collective effervescence. Indeed, PES partially mediated the relationship between QP and remembered well-being in a simple mediation.

It was expected that PES would show higher effects on the dependent variables. The lack of direct effects of PES on them may be explained because the brief scale only measures a small part of the original construct (Wlodarczyk *et al.*, 2020). It's possible, too, that the proximal effects overlap the effects of the PES because of the strong relationships between them, or that its effects act differently. Despite this, the proposed model has shown to be theoretically and empirically adequate.

Regarding the durability of the effect of participation, Durkheim underlined the need to repeat collective gatherings, since the effects dissipate over time. The few works that have analyzed this issue shows that the effects of participation in collective gatherings can last between 1 and 4 weeks (Páez *et al.*, 2007, 2011, 2015; Rimé *et al.*, 2010; Khan *et al.*, 2016; Tewary *et al.*, 2012), but even also up to 10 weeks (Bouchat, *et al.*, 2020). According to the data from our study, the effects of participation on social contribution, remembered well-being and collective empowerment dissipate at least six to seven weeks after having participated. However, the effect of participation on social integration, social acceptance and social actualization is maintained (although it decreases). Thus, the participants of Korrika came back to their daily life with a greater feeling of belonging and being accepted by their community, with a better view of people, and with a greater trust in humanity and social progress. These are important indicators of the psychological health of people (Keyes, 1998; Páez & Oyanedel, 2021; Postmes *et al.*, 2019; Zabala *et al.*, 2020) that need to be taken into account.

Another interesting outcome that deserves to be highlighted from the results is the fact that while the effects of enjoyment emotions and kama muta on these variables (i.e., social integration, social acceptance and social actualization) dissipated before the six to seven weeks after the participation, PES maintained its effect over this time. This suggests, that both enjoyment and self-transcendent emotions (kama muta included) have a powerful but fleeting effect. That is, they act tied to the moment where collective effervescence occurs (up to least 1 week according to the data of this study). Instead, the effect of PES can reach far beyond that instance, even up to 10 weeks according to the study of Bouchat *et al.* (2020). Therefore, it seems that the PES is what matters to maintain these positive

psychological effects over time. However, further investigations are needed to confirm this relatively fleeting effect of emotions and the long-lasting effect of PES.

Limitations and Future Studies

The first limitation of this study is that, despite the fact that the model presented —PES as a central mechanism of collective processes— has shown to be both theoretically and statistically more adequate than other models (see Supplementary Table S3) and in line with both the theory and the results of previous studies, we cannot confirm a causality between the PES and the proximal and distal effects for methodological reasons. However, these results provide a basis to guide future studies in which this possible causal relationship is analyzed experimentally. As for the instruments used, it would probably have been more appropriate to use the full versions, as well as instruments that more fully measure the more specific aspects of personal well-being. In addition, it is likely that the instrument used to measure the PES did not measure the construct in its entirety (Wlodarczyck *et al.*, 2020), and perhaps, for this reason, it has shown fewer effects than expected. Anyway, these instruments were selected with the aim of avoiding participant fatigue and the sample loss that commonly happen in longitudinal studies. Another limitation refers to the low reliability indices that social acceptance has shown, especially in T1, which in turn has had a negative influence on the fit indices of the social acceptance models. Therefore, this result should be viewed with caution. On the other hand, the relationships between the proximal effects of the PES and the dependent variables should also be considered with caution, as well as the possible explanations that have been presented, since these relationships could be different depending on the type of collective gathering analyzed. However, the results show a clear influence of collective effervescence on the effects of collective participation. Finally, in future studies, it would be fruitful to have a control group that had not participated in the meeting, since this would allow a better assessment of the results.

5. Conclusion

From the results of this study we can conclude, first, that participation in Korrika has positive psychological effects on social well-being and collective empowerment. In addition, the effects on social integration, social acceptance and social actualization can last for at least six to seven weeks after participation. Second, we conclude that PES can be considered one of the central mechanisms of the effects of collective participation, and of special importance for their maintenance. Third, that kama muta emerges as a relevant emotion that should be paid attention to in future studies. And, finally, given the effect that collective gatherings (and their absence) have on people's psychological health, we conclude that collective participation should be considered as another factor to be valued in health prevention programs.

6. Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7. Author Contributions

JZ, SC, and AP contributed to the design and implementation of the research. JZ collected the study's data and analyzed and interpreted the data together with LZ and JP. JZ created the first draft of the manuscript which was then translated by IA-A. JZ, SC, AP, JP, LZ, and IA-A discussed the results, commented on the manuscript and contributed to the writing and editing of the final version of this manuscript. All authors contributed to the article and approved the submitted version.

8. Funding

This research was carried out with funding provided by a Pre-doctoral Grant to JZ (PRE_2020_1_0338), by a Post-doc Grant from the UPV/EHU to LZ

(DOCBERRI 20/23), by a grant from the Basque Government for Research Groups (Ref. IT-1598-22) and by a grant from the Spanish Ministry of Science and Innovation (Ref. PID2020-115738GB-I00).

9. Data Availability Statement

The datasets generated for this study are available upon request made to the corresponding author.

10. Ethics Statement

The studies involving human participants were reviewed and approved by University of the Basque Country's Ethics Committee for Research involving Human Beings. The patients/participants provided their written informed consent to participate in this study.

11. References

Agostini, M., & van Zomeren, M. (2021). Toward a comprehensive and potentially cross-cultural model of why people engage in collective action: A quantitative research synthesis of four motivations and structural constraints. *Psychological Bulletin, 147*(7), 667-700. https://doi.org/10.1037/bul0000256

Alfaro-Beracoechea, L. N., & Contreras-Tinoco, K. A. (2021). Experiencias evocadoras de la emoción kama muta en jóvenes mexicanos: sentirse conmovido por amor [Evocative experiences of kama muta emotion in young Mexicans: feeling moved by love]. *Revista Latinoamericana de Estudios Sobre Cuerpos, Emociones y Sociedad*, *3*(37), 65-76.

Aquino, K., McFerran, B., & Laven, M. (2011). Moral identity and the experience of moral elevation in response to acts of uncommon goodness. *Journal of Personality and Social Psychology*, *100*(4), 703-718. https://doi.org/10.1037/a0022540

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Blanco, A., & Díaz, D. (2005). El bienestar social: su concepto y medición [Social well-being: its concept and measurement]. *Psicothema*, *17*(4), 582-589.

Bouchat, P., Rimé, B., Van Eycken, R., & Nils, F. (2020). The virtues of collective gatherings: A study on the positive effects of a major scouting event. *Journal of Applied Social Psychology*, *50*(3), 189-201. https://doi.org/10.1111/jasp.12649

Bueno-Notivol, J., Gracia-García, P., Olaya, B., Lasheras, I., López-Antón, R., & Santabárbara, J. (2021). Prevalence of depression during the COVID-19 outbreak: A meta-analysis of community-based studies. *International Journal of Clinical and Health Psychology*, *21*(1). https://doi.org/10.1016/j.ijchp.2020.07.007

Castro-Abril, P., da Costa, S., Navarro-Carrillo, G., Caicedo-Moreno, A., Gracia-Leiva, M., Bouchat, P., Cordero, B., Méndez, L & Páez, D. (2021). Social identity, perceived emotional synchrony, creativity, social representations, and participation in social movements: The case o the 2019 chilean populist protests. *Frontiers in Psychology*, *12*, Article 764434. https://doi.org/10.3389/fpsyg.2021.764434

Cohen-Chen, S., & Van Zomeren, M. (2018). Yes we can? Group efficacy beliefs predict collective action, but only when hope is high. *Journal of Experimental Psychology, 77*, 50-59. https://doi.org/10.1016/j.jesp.2018.03.016

Collins, R. (2004). *Interaction ritual chains*. Princeton University Press.

del Valle, T. (1988). *Korrika: Rituales de la lengua en el espacio [Korrika: Rituals of the language in the space]*. Anthropos Editorial.

Dimmock, J., Krause, A. E., Rebar, A., & Jackson, B. (2021). Relationships between social interactions, basic psychological needs, and wellbeing during the COVID-19 pandemic. *Psychology and Health*, 457-469. https://doi.org/10.1080/08870446.2021.1921178

Drury, J., & Reicher, S. (2009). Collective psychological empowerment as a model of social change: Researching crowds and power. *Journal of Social Issues*, *65*(4), 707-725. https://doi.org/https://doi.org/10.1111/j.1540-4560.2009.01622.x

Durkheim, E. (1995). *The elementary forms of religious life* (Original work published 1912). Free Press.

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689-723. https://doi.org/10.1037/0033-295X.99.4.689

Fiske, A. P., Seibt, B., & Schubert, T. W. (2017). The sudden devotion emotion: Kama muta and the cultural practices whose function is to evoke it. *Emotion Review*, *11*(1), 1-13. https://doi.org/10.1177/1754073917723167

Fredrickson, B. L. (2009). *Positivity*. Three Rivers Press.

Gómez, A., & Vázquez, A. (2015). Personal identity and social identity: Two different processes or a single one? *Revista de Psicología Social*, *30*(3), 468-480. https://doi.org/10.1080/02134748.2015.1065091

Haidt, J. (2006). *The happiness hypothesis.* Basic Books.

Hervás, G., & Vázquez, C. (2013). Construction and validation of a measure of integrative well-being in seven languages: The Pemberton Happiness Index. *Health and Quality of Life Outcomes*, *11*, Article 60. https://doi.org/10.1186/1477-7525-11-66

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, *61*(2), 121-140. https://doi.org/10.1177/0038040712456559

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*(2), 207-222. https://doi.org/10.2307/3090197

Khan, S. S., Hopkins, N., Reicher, S., Tewari, S., Srinivasan, N., & Stevenson, C. (2016). How collective participation impacts social identity: A longitudinal study from India. *Political Psychology*, *37*(3), 309-325. https://doi.org/10.1111/pops.12260

Landmann, H., & Rohmann, A. (2020). Being moved by protest: Collective efficacy beliefs and injustice appraisals enhance collective action intentions for forest protection via positive and negative emotions. *Journal of Environmental Psychology*, *71*, Article 101491. https://doi.org/10.1016/j.jenvp.2020.101491

Lizarazo Pereira, D. M., Schubert, T. W., & Roth, J. (2022). Moved by social justice: The role of kama muta in collective action toward racial equality. *Frontiers in Psychology*, *13*, Article 780615. https://doi.org/10.3389/fpsyg.2022.780615

Macdonald, B., & Hülür, G. (2021). Well-being and loneliness in swiss older adults during the COVID-19 pandemic: The role of social relationships. *The Gerontologist*, *61*(2), 240-250. https://doi.org/10.1093/geront/gnaa194

Marroquín, B., Vine, V., & Morgan, R. (2020). Mental health during the COVID-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry Research*, *293*, Article 113419. https://doi.org/10.1016/j.psychres.2020.113419

Martínez, J., Guillén, F., & Feltz, D. (2011). Psychometric properties of the Spanish version of the Collective Efficacy Questionnaire for Sports. *Psicothema*, *23*(3), 503-509.

McPhail, C., & Wohlstein, R. T. (1983). Individual and collective behaviors within gatherings, demonstrations, and riots. *Annual Review of Sociology*, *9*(1), 579-600. https://doi.org/10.1146/annurev.so.09.080183.003051

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13-20. https://doi.org/10.1016/j.jesp.2017.03.009

Novelli, D., Drury, J., Reicher, S., & Stott, C. (2013). Crowdedness mediates the effect of social identification on positive emotion in a crowd: A survey of two crowd events. *PLoS ONE*, *8*(11), 1-7. https://doi.org/10.1371/journal.pone.0078983

Páez, D., Basabe, N., Ubillos, S., & González-Castro, J. L. (2007). Social sharing, participation in demonstrations, emotional climate, and coping with collective violence after the March 11th Madrid bombings. *Journal of Social Issues*, *63*(2), 323-337.

Páez, D., Bilbao, M. A., Bobowik, M., Campos, M., & Basabe, N. (2011). ¡Feliz Navidad y próspero año nuevo! El impacto de los rituales de Navidad en el bienestar subjetivo y en el clima emocional familiar [Merry Christmas and Happy New Year! The impact of Christmas rituals on subjective well-being and family's emotional climate]. *Revista de Psicología Social*, *26*(3), 373-386. https://doi.org/10.1174/021347411797361347

Páez, D., Javaloy, F., Wlodarczyk, A., Espelt, E., & Rimé, B. (2013). El movimiento 15-M: Sus acciones como rituales, compartir social, creencias, valores y emociones [The 15-M movement: Actions as rituals, social sharing, beliefs, values and emotions]. *Revista de Psicología Social*, *28*(1), 19-33. https://doi.org/10.1174/021347413804756078

Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*(5), 711-729. https://doi.org/http://dx.doi.org/10.1037/pspi0000014

Páez, D., & Oyanedel, J. C. (2021). Social belongingness and well-being: International perspectives [Monograph]. *Frontiers in Psychology*. doi: 10.3389/978-2-88971-501-5

Pelletier, P. (2018). The pivotal role of perceived emotional synchrony in the context of terrorism: Challenges and lessons learned from the March 2016 attack in Belgium. *Journal of Applied Social Psychology*, *48*(9), 477-487. https://doi.org/10.1111/jasp.12526

Pizarro, J. J., Basabe, N., Amutio, A., Telletxea, S., Harizmendi, M., & Van Gordon, W. (2019). The mediating role of shared flow and perceived emotional synchrony on compassion for others in a mindful-dancing program. *Mindfulness*, *11*, 125-139. https://doi.org/10.1007/s12671-019-01200-z

Pizarro, J. J., Basabe, N., Fernández, I., Carrera, P., Apodaca, P., Man Ging, C. I., Cusi, O., & Páez, D. (2021). Self-transcendent emotions and their social effects: Awe, elevation and kama muta promote a human identification and motivations to help others. *Frontiers in Psychology*, *12*, Article 709859. https://doi.org/10.3389/fpsyg.2021.709859

Pizarro, J. J., Zumeta, L. N., Bouchat, P., Wlodarzyck, A., Rimé, B., Basabe, N., Amutio, A., & Páez, D. (2022). Emotional processes, collective behavior, and social movement: A meta-analytic review of collective effervescence outcomes during collective gatherings and demostrations. *Frontiers in Psychology*, *13*. Article 97468. https://doi.org/10.3389/fpsyg.2022.97468

Postmes, T., Wichmann, L. J., van Valkengoed, A. M., & van der Hoef, H. (2019). Social identification and depression: A meta-analysis. *European Journal of Social Psychology, 49*(1), 110-126. https://doi.org/10.1002/ejsp.2508

Rovai, A. P., Baker, J. D., & Ponton, M. K. (2013). *Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS*. Watertree Press LLC

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A meta-analysis. *Zeitschrift Fur Psychologie*, *224*(3), 168-189. https://doi.org/10.1027/2151-2604/a000252

Rimé, B., Páez, D., Basabe, N., & Martinéz, F. (2010). Social sharing of emotion, post-traumatic growth, and emotional climate: Follow-up of Spanish citizen's response to the collective trauma of March 11th terrorist attacks in Madrid. *European Journal of Social Psychology, 40*, 1029-1045. https://doi.org/10.1002/ejsp

Seibt, B., Schubert, T. W., Zickfeld, J. H., & Fiske, A. P. (2019). Touching the base: Heart-warming ads from the 2016 U.S. election moved viewers to partisan tears. *Cognition and Emotion*, *33*(2), 197-212. https://doi.org/10.1080/02699931.2018.1441128

Sociolinguistic cluster (2020). *Euskararen datu soziolinguistikoen biltegia* [*Sociolinguistic database on Euskera*]. Available online at: http://www. soziolinguistika.eus/edb/index.php?erakus=aurkezpena (accessed August 15).

Stellar, J. E., Gordon, A. M., Piff, P. K., Cordaro, D., Anderson, C. L., Bai, Y., Maruskin, L. A., & Keltner, D. (2017). Self-transcendent emotions and their social functions: Compassion, gratitude, and awe bind us to others through orosociality. *Emotion Review, 9*(3), 200-207. https://doi.org/10.1177/1754073916684557

Tewari, S., Khan, S., Hopkins, N., Srinivasan, N., & Reicher, S. (2012). Participation in mass gatherings can benefit well-being: Longi-

tudinal and control data from a North Indian Hindu pilgrimage event. *PLoS ONE*, *7*(10). Article e4729 https://doi.org/10.1371/journal.pone.0047291

Thonhauser, G. (2022). Towards a taxonomy of collective emotions. *Emotion Review*, *January*, 31-42. https://doi.org/10.1177/17540739211072469

Van Cappellen, P., & Rimé, B. (2014). Positive emotions and self-transcendence. In V. Saroglou (Ed.), *Religion, Personality, and Social Behavior* (pp. 123-145). Psychology Press.

Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin, 134*(4), 504-535. https://doi.org/10.1037/0033-2909.134.4.504

Wlodarczyk, A., Basabe, N., Páez, D., Reyes, C., Villagrán, L., Madariaga, C., Palacio, J., & Martínez, F. (2016). Communal coping and posttraumatic growth in a context of natural disasters in Spain, Chile, and Colombia. *Cross-Cultural Research, 50*(4), 325-355. https://doi.org/10.1177/1069397116663857

Wlodarczyk, A., Basabe, N., Páez, D., Villagrán, L., & Reyes, C. (2017b). Individual and collective posttraumatic growth in victims of natural disasters: A multidimensional perspective. *Journal of Loss and Trauma*, *22*(5), 371-384. https://doi.org/10.1080/15325024.2017.1297657

Wlodarczyk, A., Basabe, N., Páez, D., & Zumeta, L. (2017a). Hope and anger as mediators between collective action frames and participation in collective mobilization: The case of 15-M. *Journal of Social and Political Psychology*, *5*(1), 200-223. https://doi.org/10.5964/jspp.v5i1.471

Wlodarczyk, A., Zumeta, L., Basabe, N., Rimé, B., & Páez, D. (2023). Religious and secular collective gatherings, perceived emotional synchrony and self-transcendent emotions: Two longitudinal studies. *Current Psychology, 42*, 4754-4771. https://doi.org/10.1007/s12144-021-01826-0

Wlodarczyk, A., Zumeta, L., Pizarro, J. J., Bouchat, P., Hatibovic, F., Basabe, N., & Rimé, B. (2020). Perceived emotional synchrony in collective gatherings: Validation of a short scale and proposition of an integrative measure. *Frontiers in Psychology*, *11*, Article 1721. https://doi.org/10.3389/fpsyg.2020.01721

Yaden, D. B., Haidt, J., Hood, R. W., Vago, D. R., & Newberg, A. B. (2017). The varieties of self-transcendent experience. *Review of General Psychology*, *21*(2), 143-160. https://doi.org/10.1037/gpr0000102

Zabala, J., Conejero, S., Pascual, A., Alonso-arbiol, I., Amutio, A., Torres-Gómez, B., Padoan de Luca, S., & Telletxea, S. (2020). Basque ethnic identity and collective empowerment: Two key factors in well-being and community participation. *Frontiers in Psychology*, *11*. Article 606316. https://doi.org/10.3389/fpsyg.2020.606316

Zickfeld, J. H., Schubert, T. W., Seibt, B., Blomster, J. K., Arriaga, P., Basabe, N., Blaut, A., Caballero, A., Carrera, P., Dalgar, I., Ding, Y., Dumont, K., Gaulhofer, V., Gracanin, A., Gyenis, R., Hu, C., Kardum, I., Lazarevic, L. B., Mathew, L., … Alan, P. F. (2019). Kama muta: Conceptualizing and measuring the experience often labelled being moved across 19 nations and 15 languages. *Emotion*, *19*(3), 402-424. https://doi.org/10.1037/emo0000450

Zlobina, A., & Dávila, M. C. (2022). Preventive behaviours during the pandemic: The role of collective rituals, emotional synchrony, social norms and moral obligation. *British Journal of Social Psychology, 61*(4), 1332-1350. https://doi.org/10.1111/bjso.12539

Zumeta, L., Basabe, N., Wlodarczyk, A., Bobowik, M., & Páez, D. (2016a). Shared flow and positive collective gatherings. *Anales de Psicología*, *32*(3), 717-727. https://doi.org/10.6018/analesps.32.3.261651

Zumeta, L. N., Bobowik, M., Basabe, N., & Wlodarczyk, A. (2021). Participation in multicultural awareness-raising community actions: Positive effects on well-being and group efficacy. *Cultural Diversity and Ethnic Minority Psychology, 28*(3), 413-426. https://doi.org/10.1037/cdp0000426

Zumeta, L. N., Castro-Abril, P., Méndez, L., Pizarro, J. J., Włodarczyk, A., Basabe, N., Navarro-Carrilo, G., Padoan de Luca, S., da Costa, S., Alonso-Arbiol, I., Torres-Gómez, B., Cakal, H., Delfino, G., Techio, E. M., Alzugaray, C., Bilbao, M., Villagrán, L., Reyes-Valenzuela, C., Alfaro-Beracoeechea, L., … Pinto, I. R. (2020). Collective effervescence, self-transcendence, and gender differences in social well-being during 8 March demonstrations. *Frontiers in Psychology*, *11*. Article 607538 https://doi.org/10.3389/fpsyg.2020.607538

Zumeta, L. N., Oriol, X., Telletxea, S., Amutio, A., & Basabe, N. (2016b). Collective efficacy in sports and physical activities: Perceived emotional synchrony and shared flow. *Frontiers in Psychology*, *6*, Article 1960. https://doi.org/10.3389/fpsyg.2015.01960

Google Tools CC by 4.0

Hipoteka-maileguen legeak egokitu behar dira gehiegizko klausulen aurrean

Espainian, norbaitek hipoteka ordaindu ezin duenean, etxetik ateraraz daiteke hipoteka exekutatzeko prozedura baten bidez. Zordunak uste badu mailegu-kontratuaren zenbait klausula gehiegizkoak direla, auzi hori «prozedura adierazgarri» baten bidez plantea dezake epaitegietan. Hala ere, horrek ez du hipoteka betearazteko prozesua geldiarazten, eta etxea gal dezake ordaindu ezin duenak. Horregatik, Espainiako epaile batek Europako Erkidegoetako Justizia Auzitegiari eskatu zion argitzeko ea Espainiako legeak bat ote datozen Kontsumitzailea Babesteko Europako Legearekin. Legeak egokitu beharra dago gehiegizko klausulen ondorioz inork etxea gal ez dezan.

Europako araudiak azpimarratzen du profesionalek era bidezko batean tratatu behar dituztela kontsumitzaileak. Gehiegizkotzat jotzen dira kontsumitzailearen interesen aurkako desoreka handiegia sortzen duten kontratuak; bereziki, banaka negoziatzen ez direnean. Kontratu bateko klausula bat gehiegizkotzat jotzen bada, baliogabetu egiten da, baina kontratuaren gainerakoa bere horretan geratzen da. Europako direktibek, hala ere, gehiegizko klausulak erabiltzea eragozten duten mekanismoak izatera behartzen dituzte estatu kideak. Espainiako lege espezifikoek direktiba horien berri ematen dute, eta bidegabetzat jotzen dituzte kontsumitzailearen aurkako desoreka garrantzitsuak eragiten dituzten edo fede onaren printzipioak urratzen dituzten xedapenak, besteak beste.

Hipoteka-maileguek, beste kontratu batzuek ez bezala, tarte bat izan dezakete klausulak banaka negoziatzeko. Hala ere, kasu gehienetan, kreditu-etxeak ezartzen ditu baldintzak, eztabaida handirik gabe, eta mailegu-hartzaileak maileguaren zenbatekoa, interes-tasak eta ordaintzeko baldintzak hartzen ditu kontuan, batez ere. Kasu horretan, klausula batzuk sartu ziren, hala nola urteko % 18,75eko interesa ordainketetan berandutzeagatik, eta mailegu-emaileak guztia itzultzeko eskatzeko eskubidea izatea mailegu-hartzaileak ordaintzen ez badu. Europako Justizia Auzitegiak ez ditu berariaz kalifikatzen klausula horiek, baizik eta jarraibideak ematen ditu estatuko epaileek kasuz kasu beren inpartzialtasuna zehaztu dezaten.

Noiz den klausula bat egokia

Klausula baten ekitatea epaitzeko, mailegu-emailearen eta mailegu-hartzailearen arteko eskubideen eta betebeharren oreka ebaluatu zuten, eta, elkarren arteko adostasunik ez badago, estatuko legeria hartzen da kontuan. Epaileak zehaztu behar du kontsumitzaile batek klausula bat onartu ote zuen, baldin eta banaka negoziatu bazen, eta mailegu-emaileak fede onez jokatu zuela bermatu. Berandutza-interesak direnean, epaileak berandutza-interesen eta legezko interes-tasaren arteko aldea aztertuko du. Zor bat bere kabuz likidatzen bada, funtsezkoa da bermatzea kontsumitzailea justiziara erraz iristeko moduan dagoela bere eskubideak defendatzeko.

Europar Batasunak 2013ko martxoaren 14an eman zuen epaia. Horren ondorioz, legea aldatu zen. Bereziki, 1/2013 Legea, maiatzaren 14koa, zorraren berregituraketa, alokairu soziala eta hipoteka zordunen babesa indartzekoa. Lege horrek, hipoteka-maileguen kontratuak izan ditzakeen klausula batzuk arautzeaz gain, abusuzko klausulen alegazioa barneratzen du hipotekaren exekuzio-prozedura eteteko arrazoi gisa, bertan jasotzen diren baldintzak betez. Aurrerantzean ere legeak bermatu beharko du inork etxea gal ez dezala abusuzko klausulen ondorioz.

Hipoteka maileguak eta kontsumitzaileen babesa: kontratuen ejekuzioa eta abusuzko klausulak (EBJAren -1. Aretoa- 2013ko martxoaren 14ko epaia)

Itziar Villafañez Perez

Euskal Herriko Unibertsitatea

LABURPENA: Hipoteka maileguak ez ordaintzeagatik herritarrak haien etxetik kaleratzeko prozedura kopuru izugarria Espainiar Estatuko gizarteak bizi duen egoeran, Europar Batasuneko Justizia Auzitegiaren 1. Aretoak, iragan martxoaren 14an, Espainiar Estatuaren Zuzenbidean, eta bereziki epaitegien jardunean, izugarrizko garrantzia duen epaia eman zuen. Honen arabera, Espainiako auzibide eta hipoteka legedia kontsumitzaileen babeserako Europar Zuzenbidearen kontrakoak dira, hipotekaren exekuzioaren erraztea bilatuz, abusuzko klausulen aurrean kontsumitzaileen interesen defentsa oztopatu edo ezinezko egiten dutelako.

1. Sarrera

Ordainketa faltagatik hipotekaren exekuziorako prozeduraren ondorioz, zorduna bere etxetik kanporatua da. Kaleratzea baino lehenago zordunak epaitegietan prozedura adierazlea hasteko demanda aurkezten du, mailegu kontratuaren klausula batzuk baliogabetzeko eskatuz, hauek abusuzkoak direlakoan, eta, ondorioz, exekuzio prozeduraren baliogabetasuna eskatuz.

Espainiako Zuzenbidearen arabera, abusuzko klausulen azterketa hipotekaren exekuzio prozedura ez den prozedura adierazlean egin beharko da, arrazoi honengatik hipotekaren exekuzio prozedura eten ezin daitekeelarik. Hau dela eta, eta Hipoteka legedia ere kontuan izanda, kontsumitzailearen etxebizitza hipoteka exekuzio prozeduraren ondorioz besterenganatu ahal izango da, normalean atzeraezina izango dena, hipoteka maileguaren klausula batzuk abusuzkoak direla eta exekuzio prozeduraren baliogabetasuna adierazten denean ere.

Egoera honen aurrean, prozedura adierazleaz ezagutzen duen epaileak Europar Batasuneko Justizia Auzitegiari Espainiar Zuzenbidea kontsumi-

* **Corresponding author / Harremanetan jartzeko:** Itziar Villafañez Perez. Zuzenbide Fakultatea, Enpresaren Zuzenbidea saila, Manuel Lardizabal 2, 20018 Donostia. e-mail: itziar.villafanez@ehu.es. **Contributed by** Aitor Zurimendi **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Villafañez Perez, Itziar. «Hipoteka maileguak eta kontsumitzaileen babesa: kontratuen ejekuzioa eta abusuzko klausulak (EBJAren -1. Aretoa- 2013ko martxoaren 14ko epaia)» https://www.eusko-ikaskuntza.eus/es/publicaciones/hipoteka-maileguak-eta-kontsumitzaileen-babesa-kontratuen-ejekuzioa-eta-abusuzko-klausulak-ebjaren-1-aretoa-2013ko/art-23521/ Eleria: Euskal Herriko legelarien aldizkaria, ISSN 1137-1951, N.º 24, 2013, págs. 71-75

tzaileen babeserako Europar Zuzenbidearekin bat datorren galdetzen dio, aldi berean hipoteka maileguan jasotzen diren klausula batzuen baliozkotasunari buruz ere galdetuz[1].

2. Oinarria: Kontsumitzaileen babeserako arauketa Europar Batasunean eta Espainian. Bereziki, kontsumitzaileekin egindako kontratuetako abusuzko klausulei buruzko arauketa

Kontseiluaren 93/13/CEE Zuzentarauak, 1993ko apirilaren 5ekoa, kontsumitzaileekin egindako kontratuetako abusuzko klausulei buruzkoa, oinarri bezala du profesionalek fede onez jardun behar dutela, beste aldea (kontsumitzailea) leialtasunez eta berdintasunez tratatuz, eta honen interes zilegiak kontuan izanda[2].

Aldi berean, bere 3. artikuluak ezartzen du banaka negoziatu ez diren kontratuetako klausulak (bereziki atxikimendu-kontratuen kasuan) abusuzkoak izango direla, fede onaren baldintzak bete arren, kontratutik alderdientzako sortutako eskubide eta betebeharren artean desoreka garrantzitsua dagoenean kontsumitzailearen interesen kontra.

Desoreka edo abusu hau existitzen den baloratzeko irizpide ezberdinak izan beharko dira kontuan, hala nola kontratuaren objektu diren ondasun eta zerbitzuak, kontratua egin zeneko inguruabar guztiak, kontratuaren (edo lotuta dauden kontratuen) gainontzeko klausulak (4. artikulua).

Zuzentarauak berak bere gehigarrian abusuzko klausulen adibideak jasotzen ditu, haien artean bere betebeharrak betetzen ez dituen kontsumitzaileari indemnizazio altuegia ezartzea, edo kontsumitzaileari akzio judizialak edo errekurtsoak erabiltzea zailtzea edo horretarako aukera kentzea (esaterako bere esku dauden frogabideak behar ez bezala mugatuz, edo legearen arabera beste alderdiari dagokion frogaren zama ezarriz).

Klausula bat abusuzkoa denean, ondorioa honen baliogabetasuna izango da, baina kontratuak klausula horiek gabe biziraun badezake, kontratua termino berdinetan bete beharrezkoa izaten jarraituko du alderdientzat (6. artikulua).

Aldi berean, Zuzentarauak Estatuei agintzen die, kontsumitzaileen eta lehiakide profesionalen intereserako, profesional eta kontsumitzaileen arteko kontratuetako abusuzko klausulen erabilpenarekin bukatzeko bide egoki eta eraginkorrak existitzen direla zaintzea (7. artikulua).

Espainiar estatuari dagokionez, gaur egun abusuzko klausulekiko kontsumitzaileen babesa azaroaren 16ko 1/2007 Erret Dekretu Legegilean, Kontsumitzaile eta Erabiltzaileen Babeserako Lege Orokor eta bestelako lege osagarrien testu bateratua onartzen duena, dago jasota.

Honen 82. artikuluaren arabera, eta Zuzentarauaren aginduak jarraituz, abusuzkoak dira banaka negoziatu ez diren xedapenak eta espresuki onartu ez diren praktikak, fede onaren eskakizunen kontra kontsumitzaile eta erabiltzaileen kalterako kontratutik eratorritako betebehar eta eskubideen artean desoreka garrantzitsuak sortzen dituztenean. Abusuzko izaera baloratzeko ondasun eta eskubideen izaera, kontratua burutu zeneko inguruabar guztiak eta kontratuaren (edo lotuta dauden kontratuen) gainontzeko klausulak izango dira kontuan.

Artikulu beraren arabera klausulak abusuzkoak izango dira a) kontratua enpresariaren borondatearekin lotzen dutenean, b) kontsumitzaile edo erabiltzailearen eskubideak mugatzen dituztenean, c) elkarrekikotasun falta suposatzen dutenean, d) kontsumitzaile edo erabiltzaileari gehiegizko bermeak edo modu desegokian frogaren zama inposatzen dizkiotenean, e) kontratuaren perfekzionatze edo exekuzioarekiko desproportzionatuak direnean, edo f) aplikagarriak diren Zuzenbide eta konpetentziari buruzko arauen kontrakoak direnean.

[1] Oso modu laburrean bada ere, epai honen erreferentzia PIZARRO MAQUEDA, María José, «Pronunciamiento del TJUE sobre la normativa española en materia de ejecución hipotecaria». In: *Revista Aranzadi Doctrinal*, 1. zkia., 2013; 213-215 orr. etan topa dezakegu.

[2] Zuzentarau honen eta bere edukiaren inguruan, ikusi esaterako PAGADOR LÓPEZ, Javier, *La directiva comunitaria sobre cláusulas contractuales abusivas*, Madril: Marcial Pons, 1998; bereziki 15-50 orr. artikulu honi dagokionez.

3. Hipotekaren exekuzio prozedura kontsumitzaileen eskubideen mugapen bezala

Epaian aztertzen den aspektu nagusia Espainiako legediak Estatuak abusuzko klausula erabilpenarekin bukatzea zailtzen duen da.

Alde batetik, Prozedura Zibilaren Legeak Hipotekaren exekuzio prozedura arautzean, honelako prozedura batean exekutatuak erabil dezakeen oposiziorako eskubidea modu garrantzitsuan mugatzen du, izan ere hurrengo kasuetan besterik ez da onartuko (eta ondorioz, kasu hauetan besterik ez da hipotekaren exekuzioa etengo): a) Bermea edo bermatutako betebeharra iraungitzea, b) Eska daitekeen zenbatekoa zehazterakoan akatsa, bermatutako zorra exekutatzaile eta exekutatuaren arteko kontu bat ixtearen saldoa bada, eta c) prozedura hastea eragin duen zama baino lehenagoko bahia edo hipoteka existitzea (695. artikulua).

Lege beraren 698 artikuluaren arabera, zordunak edo hirugarren batek formulatu dezakeen beste edozein erreklamazio, legean jasotakoen artean ez badago (funtsean, aurreko zioez gain, auzubidezkoarrazoi oinarritzen direnak), dagokion epaiketan konponduko da, inolaz ere hipotekaren exekuzio prozedura eten edo oztopatu dezakeelarik, nahiz eta erreklamazioa tituluaren baliogabetasunaren, edo zorraren epea, ziurtasuna, iraungitzea edo zenbatekoaren ingurukoa izan. Hala ere, posible da bigarren epaiketa honen balizko epaiaren eraginkortasuna bermatzeko hartzekodunari eman beharreko kopurua edo honen zati bat atxikitzea eskatzea. Honen guztiaren oinarria hipotekaren eta kontratutik eratorritako betebeharren arteko ezberdintasunean topatu beharko genuke, lehenengoaren bigarren mailako edo erantsi izaera nolabait albo batera utziz, eta hartzekodunari pribilegio izugarria emanaz[3].

Aurrekoaren arabera, exekutatuak azaltzen duen bezala, kreditu entitateak balizko zorra kobratzeko hipotekaren exekuzio prozedura hautatzekotan, mailegu kontratuak izan ditzakeen klausulen abusuzko izaera alegatzeko aukerak oso mugatuak dira, izan ere hauek ondorengoko prozedura adierazle batean argitu beharko dira, hipotekaren exekuzioa etengo ez dutelarik. Hau da, hipoteka horretarako prozedura bereziaren bitartez berehala exekutatuko den bitartean, honi lotutako mailegu kontratuaren klausulen baliozkotasuna prozedura adierazle batean argituko da, hipoteka dagoeneko exekutatua izan denean eta kontsumitzaileari etxebizitza kendu zaionean, kontsumitzaileari babes eraginkorra bermatzea zeharo zailduz.

Gainera Hipoteka Legearen 131 artikuluaren arabera hipotekaren baliogabetasunaren demandaren edo exekuzioa eten dezaketen aipatutako arrazoietan oinarritzen ez direnen oharpen prebentiboak ezeztatuko dira, hauek zamen ziurtagiria emateaaren alboko oharraren ostekoak badira. Honela, Espainiako auzibidezko sistema arabera hirugarren bati hipotekatutako ondasuna adjudikatzen zaionean, hau atzeraezina izango da, nahiz eta kontsumitzaileak aurkaratutako klausularen abusuzko izaerak hipotekaren exekuzio prozeduraren baliogabetasuna suposatu, salbu kontsumitzaileak demandaren oharpen prebentiboa aipatutako alboko oharra baino lehenago egiten duenean.

Arrazoi hau dela eta, arauketa hau Zuzentarauaren aginduen kontra doa, izan ere hipotekaren exekuzio prozeduran ez da posible klausulen abusuzko izaeran oinarritutako oposizio zioak formulatzea, eta aldi berean klausulen abusuzko izaeraren inguruko prozesu adierazlea ezagutzen duen epaileak ezin du bere azken erabakiaren eraginkortasun osoa bermatuko duen erresalbuzko neurririk hartu.

Zuzentarauak kontsumitzaileei ematen dien babesaren oinarria hauek enpresari edo profesionalarekiko dauden desabantaila egoeran dago. Honen ondorioz, abusuzko klausulek ez dute kontsumitzailea lotuko, eta Zuzentarauaren aplikazio esparruaren barruan epaileek klausulen abusuzko izaera ofizioz hauteman beharko dute, kontsumitzaile eta profesionalaren eskubide eta betebeharren arteko oreka eta berdintasuna bilatuz.

[3] Sakonago MONTERO AROCA, Juan, *Ejecución de la hipoteca inmobiliaria*, Valencia: Tirant lo Blanch, 2012: 1103-1292 orr. (bereziki 1179-1292 orr.); ADAN DOMÈNECH, Federic, *La ejecución hipotecaria*, Barcelona; Bosch, 2009: 399-443 orr.

Hipotekaren exekuzio prozedura, honen kontrako oposizio zioak eta kontratuaren klausulen zilegitasuna aztertzeko prozedura adierazleaz ezagutzen duen epailearen ahalmenak printzipioz Estatu bakoitzaren eskumenen barruan daude, Estatuen auzibidezko autonomiaren printzipioaren arabera, baldin eta barne Zuzenbidean antzeko egoerekin alderatuz kontrakoagoak ez direnean (ekibalentzia printzipioa), eta Europar Batasuneko ordenamendu juridikoak kontsumitzaileei aitortzen dizkien eskubideak egikaritzea zailegi edo ezinezko egiten ez duen bitartean (eraginkortasun printzipioa).

Ekibalentzia printzipioari dagokionez, kasu honetan errespetatzen da, izan ere hipotekaren exekuzio prozedurari lotutako prozedura adierazleaz ezagutzen duen epaileak beste kasu batzuetan ere ez du gaitasunik izango bere azken erabakiaren eraginkortasuna bermatuko duten badaezpadako neurriak hartzeko, esaterako ordena publikoaren klausulen kasuan.

Aldiz, eraginkortasun printzipioari dagokionez, arlo honetan Espainiako Zuzenbidea Zuzentarauaren aginduen kontrakoa da. Aipatu bezala, Espainiako legediaren ondorioz, hipotekaren baliogabetasunaren demandaren edo exekuzioa eten dezaketen aipatutako arrazoietan oinarritzen ez direnen oharpen prebentiboak ezeztatuko dira, hauek zamen ziurtagiria ematearen alboko oharraren ostekoak badira, eta honen ondorioz normalean hipotekatutako ondasuna besterenganatu ahal izango da, eta nahiz eta ondorengoko epai batek klausulen abusuzko izaera adierazi eta hipoteka exekutatzeko prozeduraren baliogabetasuna suposatu, besterentze hori atzeraezinezkoa izango da. Beraz, prozedura adierazlea ezagutzen duen epaileari bere azken erabakiaren eraginkortasuna bermatzeko badaezpadako neurriak (bereziki hipotekaren exekuzioa etetea) hartzea galarazten zaion heinean, Zuzentarauak kontsumitzaileei eman nahi dien babesaren eraginkortasuna urrituko da, dagoeneko hipoteka exekutatu denean kontsumitzaileari *a posteriori* kalte-ordain bat besterik ez baitzaio emango klausulak abusuzkoak direla adierazten denean, argi eta garbi bere interesak defenditzeko nahikoa ez dena, bereziki hipotekatutako ondasuna kontsumitzailearen eta bere familiaren etxebizitza denean.

4. Abusuzko klausulak hipoteka mailegu kontratuetan

Aldi berean, EBJAren epaiak hipoteka mailegu kontratuak jaso ditzakeen klausula batzuk aztertzen ditu, abusuzkoak izan daitezkeen erabakitzeko.

Kontuan izan behar da hipoteka maileguetan beste kontratu batzuetan baino toki gehiago izan dezakeela klausulen banakako negoziazioak. Hala ere, normalean negozioa aspektu oso konkretuetara mugatzen da, kontsumitzaileak bere interesa jaso behar duen zenbatekoan, ordaindu beharreko ohiko interesetan eta komisioetan eta epemugan zentratzen baitu[4]. Beraz, klausula gehienak kreditu entitateak barneratzen ditu kontratuan inolako negoziaziorik gabe, kontratuaren baldintza orokorren parte izanik.

Beste batzuen artean, aztertzen den kasuan honako klausulak jasotzen zituen kontratuak: a) Urteroko %18,75ko berandutze interesak, epemugan ordaindu ez diren zenbatekoekiko zuzenean aplikagarriak, inolako erreklamaziorik egin beharrik gabe; b) Adostutako epemugetariko batean zordunak kapitalaren zati bat edo maileguaren interesak ez ordaintzekotan, kreditu entitateak maileguaren zenbateko osoa eskagarria dela adierazteko gaitasuna (hau da, aurretiazko epemugaketa), eta c) kreditu entitateak zor posiblea kobratzeko hipotekaren exekuzio prozedurara jotzeko aukera, eta honetarako zuzenean eska daitekeen zenbatekoa likidatzeko gaitasuna hau jasotzen duen ziurtagiriaren bidez («likidezia paktua»).

Zentzu honetan, EBJAk ez du argitzen klausula hauek zuzenean abusuzkoak diren ala ez, horren ordez epaileek kasuz kasu erabakia hartu dezaten irizpide batzuk ezarriz, Zuzentarauaren 3. Artikuluak jasotako ideia orokorretatik abiatuz[5].

[4] ANGUITA RÍOS, Rosa M., *Constitución y ejecución del crédito hipotecario*, Madril: Marcial Pons, 2008, 37-38. orr.

[5] Esan beharra dago doktrinak eta jurisprudentziak klausula hauen (eta orokorrean hipoteka maileguen kontratuetan dauden klausulen) abusuzko izaera hainbattetan azteru dituztela. Aipagarriak dira zentzu honetan ANGUITA RÍOS, Rosa M., *Constitución y ejecución(…)*, 49-63 orr.; MARTÍNEZ DE SALAZAR BASCUÑANA, Lucio, *Condiciones generales y cláusulas abusivas en los contratos bancarios*, Cadiz; EDICIP: 2002: 280-288 eta 306-308 orr. Orokorrean ere De TORRES PEREA, José Manuel,

Honela, alderdien eskubide eta betebeharren arteko «desoreka garrantzitsua» dagoen aztertzeko, alderdien akordiorik ez dagoenerako barne Zuzenbidearen arabera aplikagarriak diren arauak izan behar dira kontuan. Aldi berean, kontsumitzaileak abusuzko klausulen erabilpena bukarazteko dituen baliabideak izan beharko dira ere kontuan.

«Fede ona»ri dagokionez, epaileak profesionalak, kontsumitzailea leialtasun eta berdintasunez tratatu izan balu, azken honek (kontsumitzaileak) klausula hori banaka negoziatuz onartu izango lukeela estimatuko lukeen aztertu beharko du.

Gainera, esan bezala, Zuzentarauaren 4. artikuluaren arabera kontratuaren objektu diren ondasun eta zerbitzuen izaera izango da kontuan, kontratua burutzeko momentuko inguruabar guztiak kontuan izanda.

Honela, aurretiazko epemugaketaren kasuan, epaileak aztertu beharko du kontsumitzaileak kontratuaren arabera esentziala den betebeharren bat bete ez duen, ez-betetzearen larritasuna maileguaren zenbatekoa eta iraupena kontuan izanda, profesionalaren fakultate hau gai honi aplikagarri zaizkion arauei dagokienez salbuespen bat den, eta barne Zuzenbideak kontsumitzaileak maileguaren aurretiazko epemugaketaren ondorioekin bukatzeko baliabide egoki eta eraginkorrak jasotzen dituen.

Berandutze interesei dagokienez, baloratu beharko dira alderdien akordiorik egon ezean barne Zuzenbidearen arabera aplikagarri liratekeen arauak, eta berandutze interesaren eta legezko interesaren arteko ezberdintasuna, interes mota hauen helburuak zeintzuk diren kontuan izanda eta hauek betetzeko beharrezkoa denaren gainetik ez daudela konprobatuz.

Azkenik, zorraren alde bakarreko likidazioaren kasuan, berriro ere kontuan izan beharko dira alderdien arteko akordiorik egon ezean aplikagarriak diren arauak, eta ea, dituen auzibidezko baliabideak aintzat hartuz, kontsumitzaileari justiziara jotzea eta bere eskubideak defenditzea zailtzen zaion.

5. Ondorengo lege aldaketei erreferentzia

Epai honen azterketatik (eta beraz artikuluaren edukitik) kanpo geratzen bada ere, ezin dugu artikulu hau bukatu epaiaren osteko legearen aldaketak aipatu gabe, bereziki 1/2013 Legea, maiatzaren 14koa, zorraren berregituraketa, alokairu soziala eta hipoteka zordunen babesa indartzekoa.

Aztertutako aspektuei dagokienez, eta egun gizartean eztabaidatzen diren beste batzuk alde batera utzita (esaterako etxebizitza ordainketan emateko aukera), Lege honek, hipoteka mailegu kontratuak izan ditzakeen klausula batzuk minimoki arautzeaz gain (esaterako, berandutze interesak mugatu egiten ditu legezko interesaren hirukoitzera, edo aurretiazko epemugaketarako gutxienez 3 hilabetetan ordaintzeko betebeharra ez betetzea beharrezkoa izango da), abusuzko klausulen alegazioa barneratzen du hipotekaren exekuzio prozedura eteteko arrazoi bezala, bertan jasotzen diren baldintzak betez.

6. Bibliografia

ADAN DOMÈNECH, Federic, *La ejecución hipotecaria*, Barcelona; Bosch, 2009.

ANGUITA RÍOS, Rosa M., *Constitución y ejecución del crédito hipotecario*, Madril: Marcial Pons, 2008.

MARTÍNEZ DE SALAZAR BASCUÑANA, Lucio, *Condiciones generales y cláusulas abusivas en los contratos bancarios*, Cadiz; EDICIP: 2002.

MONTERO AROCA, Juan, *Ejecución de la hipoteca inmobiliaria*, Valencia: Tirant lo Blanch, 2012.

PAGADOR LÓPEZ, Javier, *La directiva comunitaria sobre cláusulas contractuales abusivas*, Madril: Marcial Pons, 1998.

PIZARRO MAQUEDA, María José, «Pronunciamiento del TJUE sobre la normativa española en materia de ejecución hipotecaria». In: *Revista Aranzadi Doctrinal*, 1. zkia., 2013; 213-215. orr.

De TORRES PEREA, José Manuel, «Capítulo XII: Cláusulas abusivas en la contratación de préstamos con garantía hipotecaria para la financiación de compra de vivienda». In: *Cláusulas abusivas en la contratación inmobiliaria*, Madril: Tecnos, 2006; 233-265 orr.

«Capítulo XII: Cláusulas abusivas en la contratación de préstamos con garantía hipotecaria para la financiación de compra de vivienda». In: *Cláusulas abusivas en la contratación inmobiliaria*, Madril: Tecnos, 2006: *passim*.

Etorkin latinoamerikarrek «ongizatearen paradoxa» pairatzen dute

Latinoamerikatik iritsitako emakumeak paradoxa konplexu batean bizi dira Espainian: jatorrizko herrialdean geratu diren senideak ekonomikoki mantentzen dituzte, haien egoera ekonomiko eta soziala hobetzeko, baina beraiek Espainiako ongizate sozialeko oinarrizko zerbitzuetatik kanpo geratzen dira maiz. Beharrezkoa da haien lan-baldintzak hobetzea, zaintza-sistema publikoa indartzea eta zaintzaile profesionalak kontratatzeko diru-partiden egokitasuna berrikustea.

Europar Batasunean bizi den biztanleria migratzaile latinoamerikarraren % 49,1 Espainian finkatu da. Horretan eragin zuzena izan dute Espainiako biztanleria zahartzeak eta adinekoak zaintzeko sistema publikoaren gabeziek, gero eta etxeko langile gehiago behar baitira. Bestetik, migratzaile latinoamerikarrei besteei baino migrazio-politika laxoagoa ezarri izanak ere ekarri du gehienak Espainian finkatzea. 2020an, 454.000 etxeko langileetatik % 64 ziren emakume migratzaileak.

Aipatutako emakume latinoamerikarrentzat, Espainian egiten duten zaintza-lan profesionala ezinbestekoa da jatorrizko herrialdean geratu den familia-egitura ekonomikoki mantentzeko ere. Belaunaldi arteko ezkutuko akordioek behartu egiten dituzte guraso, aiton-amona, seme-alaba eta bestelako senideei dirua bidaltzera. Askotan, gainera, gerora ere, Espainian familia berria sortuta ere, ezkutuan egiten dute, fideltasunez.

«Etorkin latinoamerikarrek jatorrizko familiari eskaintzen dioten ongizatea nekez lortzen dute eurentzat Espainian»

Beraz, belaunaldien arteko akordio informal horien eta gero gerta daitezkeen bikote-hausturen arteko orekak baldintzatzen ditu emakume horien estatuz gaindiko babeserako estrategiak. Aldi berean, urrutiko amatasunak kostu emozional handia du emakume migratzaileentzat, jatorrizko herrialdean geratu diren seme-alabei behar duten arreta eta maitasuna eman ezin izateagatik. Horrek guztiak zaurgarritasun handiko egoera batean jartzen ditu emakume migratzaileak.

Alabaina, paradoxa betean bizi dira: jatorrizko familiari eskaintzen dioten ongizatea nekez lortzen dute eurentzat Espainian. Izan ere, Espainian lan-egonkortasunak eta erresidentzia-estatutuak ematen dute zerbitzu sozialetarako eskubidea; beraz, emakume migratzaile zaintzaileek ez dute izaten babes sozial formalik. Politika honek ez du ondo egituratzen ongizatearen hartu-emana.

Politika publikoen albo-ondorioak

Ikertzaileen ustez, politika publiko espainiarren eragin zuzena da “ongizatearen paradoxa”. Batetik, etxeko langileak kontratatzea sustatzen du, familiei laguntza ekonomikoak ematen baitizkie etxeko langileak har ditzaten, zaintza-sistema publikoa indartu beharrean. Eta, era berean, lan-araudiak blokeatu egiten du etxeko langile migratzaile horien oinarrizko zerbitzu sozialetarako eskubidea. Egiturazko faktoreek baldintzatzen dute haien ongizatea.

Beraz, ikertzaileek diote berraztertu egin behar dela egokiak ote diren etxekoen zaintzarako ematen diren diru-partidak, legedia-testuinguru horretan eragiten dituzten albo-ondorioak direla eta. Mantentzekotan, emakume migratzaile horien guztien baldintzak hobetzeko neurriak hartu beharko dira. Ezinbestekoa ikusten dute, halaber, zaintza-sistema publiko indartsu bat izateko neurri politikoak hartzea.

InpAkta 00 (2023) 231-244

https://doi.org/10.17645/si.v10i1.4639

Welfare paradoxes and interpersonal pacts: transnational and gendered strategies for social protection of Latin American migrants in Spain

Laura Oso, Raquel Martínez-Buján

Department of Sociology and Communication Sciences, University of A Coruña

ABSTRACT: This article analyses the relationship between migration, care work and welfare provision, highlighting the role of Latin American migrants in Spain as providers of formal and informal social protection on a transnational scale. It contributes to the debate on "transnational social protection" and "transnational social inequalities" from the perspective of "welfare paradoxes" and "interpersonal pacts". Migrant women in Spain have become a resource for the provision of formal social protection through their employment as domestic care workers. Nevertheless, given that access to social rights in Spain depends on job stability and residency status, they have difficulties in accessing formal social protection themselves. This process constitutes a "welfare paradox", based on "commodification and exclusion paradoxes", explained by structural factors such as the characteristics of the welfare regime (a family model with privatization based on hiring migrant caregivers); the migration regime (feminised and with a clear leaning towards Latin American women); and the economic landscape resulting from two systemic crises: the Great Recession of 2008 and the COVID-19 crisis. Interpersonal pacts, rooted in marriage/couple and intergenerational agreements, and their infringements, are analysed to explain the transnational and informal social protection strategies in a context of the "exclusion paradox" and the breach of the "welfare pact". The research draws on the exploitation of secondary data and a multi-sited, longitudinal fieldwork, based on biographical interviews conducted with various members of transnational families in Spain and Ecuador (41 interviews).

1. Introduction

Scholarship studying the link between migration, care work and welfare provision has traditionally focused on the drivers of "global care chains" (Hochschild, 2000), "social care" analysis (Daly & Lewis, 2000) and the "circulation of care" (Baldassar & Merla, 2014). In recent years, "transnational social protection" studies (hereafter TSP) have reviewed reflections on these analyses and contributed new approaches to them.

* **Corresponding author / Harremanetan jartzeko:** Laura Oso. Department of Sociology and Communication Sciences, University of A Coruña, Spain. e-mail: laura.oso@udc.es. **Contributed by** Felix Arrieta Frutos **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Oso, Laura *et al.* «Welfare Paradoxes and Interpersonal Pacts: Transnational and Gendered Strategies for Social Protection of Latin American Migrants in Spain», ISSN: 2183–2803. 2022, Volume 10, Issue 1, Pages 194.204 (https://doi.org/10.17645/si.v10i1.4639)

Recently, this analytical framework has been used to argue that the "global care chain" concept was focused principally on highlighting the emotional costs of distance mothering paid by migrant women employed as caregivers and unable to provide their children with the necessary attention and affection. Likewise, the "social care theory" has provide extremely useful in analysing the social models of care and their evolution, despite the fact that they are rooted in the notion of social rights associated with nation-states, thereby complicating the identification of transnational protection needs (Parella & Speroni, 2018). Furthermore, it assimilates social protection with the formal sphere, overlooking the informal protection practices that are forged in personal interaction networks. Finally, and although the "circulation of care" concept creates broader areas of care than the previous concepts, it is unable to link the demand for care with "the precarious labour market conditions for migrant care givers" (Lutz, 2018, p. 582), and therefore fails to articulate the exchange of welfare with the structure of global inequality that is the natural habitat of these workers.

In an attempt to bridge these gaps, TSP research has introduced new conceptualisations that have contributed to an understanding of how the global inequality processes of migrant families (in particular those of female migrant domestic care workers) are articulated with their own transnational care strategies. As example of this is the term "assemblage of care", coined by Amelina (2020, 2017). This author argues that the concept allows care to be distanced from the territorial category in which its provision is articulated, allowing for the identification of patterns of inequality that emerge precisely during this assemblage. A further contribution to this process is the use of the term "transnational social inequality" (Lutz & Amelina, 2019), which refers to the way in which the transnational nature of care implies the creation of new hierarchical patterns, which naturally stem from gendered and racialised cultural values regarding care provision, but also from the support for their organisation proffered by public regulations in the host countries, the economic situation and the regulations governing domestic service. This concept, studied mainly in relation to migrant domestic care workers, links the disadvantaged position of these workers in both the labour market and in the social structure of the host country. This inequality has a localised impact, reflected in migrant marginalisation in relation to the public schemes of social protection, but it also operates in a transnational space as it generates the deployment of cross-border support practices within the family in order to overcome situations of social risk.

Rooted in the intention to move forward with the theoretical and empirical articulation between the structural processes conditioning global inequalities in access to welfare and the informal practices of mutual support and solidarity deployed among transnational families, the principal objective of this article is to apply the concepts of "welfare paradoxes" and "interpersonal pacts" (focusing on intergenerational and marriage/couple pacts) to the analysis of the formal and informal strategies for the transnational social protection of Ecuadorian migrants in Spain. This article contributes to scholarship in this field firstly through its application of the concept "welfare paradox" to the study of the adverse effects of welfare policies and their connection with cross-border inequalities resulting specifically from formal welfare provision. The "welfare paradox" concept was initially used to analyse the redistribution capacity of universalist welfare policies (Korpi & Palme, 1998), and more recently to assess the gender-equality impact of work-family reconciliation policies (Kowalewska, 2021). In this reading, we use the "welfare paradox" to consider the impact of Spanish long-term care policies, based on cash transfers, on the demand for domestic migrant and care workers, as well as the conditioning labour factors that block these workers' access to basic social rights. We argue that this "welfare paradox" is supported by a further two paradoxes, that are conceptualised and discussed here, namely the "commodification paradox" and the "exclusion paradox".

Secondly, the article contributes to scientific production debates, using the concepts of "intergenerational and couple/marriage pacts" to explore informal TSP strategies. We have defined the concept of "intergenerational pact" as an implicit agreement, rooted in cultural norms underlying the family sense of solidarity regarding the provision of care between generations (Ayuso, 2012). Beyond the formal marriage contract,

which implies a series of rights and obligations for the couple regulated by law, we consider marriage/couple pacts as a series of informal agreements based on their relationship and which organise the provision of family welfare. These pacts are obstructed by cultural norms and gender impositions, which traditionally reserve reproductive work for women, and confer the traditional role of "breadwinner" on men. The article highlights the way in which intergenerational and marriage pacts account for the articulation of transnational strategies of informal social protection. Such strategies are used by transnational families to offset the effects of the "welfare paradoxes" and their exclusion from the social pacts that are implicit in the nature of welfare states. The originality of our article lies not merely in the fact that it analyses mutual assistance and family support strategies, which have been amply addressed in literature, but also because it highlights the manner in which conflicts and the breakdown of "intergenerational and marriage/couple pacts" account for the explanation of the formal and informal social protection strategies deployed in the transnational space.

Thirdly, the article contextualises transnational social protection from a top-down approach, considering informal assistance practices, Spanish welfare policies and the economic landscape. Previous analyses have been limited to understanding the exchange of goods and services in the regions where migrants settle, whilst structural factors shaping these practices have been overlooked (Hellgren & Serrano, 2017). "Welfare paradoxes" and "interpersonal pacts" are essentially conceptual tools that connect responsibilities for care (both those held with the transnational family and those that emerge as a result of entering into paid care work) with the economic and political situation of the host country, whilst also revealing how these scenarios activate specific transnational welfare practices.

The paper is structured as follows: the next section describes the methodology used, and is followed by a review of the manner in which Latin American migrant women have become the main providers of "formal social protection" in Spain, addressing in particular social policies aimed at long-term care and describing the "welfare paradox". The following section broadens this examination by detailing the informal social protection strategies deployed by Ecuadorian transnational families through the application of the concept of "intergenerational and marriage/couple pacts". The final section presents the principal conclusions.

2. Methodology

The methodology is based on the analysis of secondary data to consider the position of Latin American migrant women within the Spanish care model and to explore the scope of formal social protection. The data were obtained from the Spanish Labour Force Survey (hereafter EPA), the Municipal Residents' Census and social service statistics from the Spanish Institute for the Elderly and Social Services (IMSERSO in its Spanish initials) (Martínez-Buján, dir., 2021-2024)

Secondly, the analysis of transnational family social protection strategies is based on multi-sited fieldwork conducted in Madrid and Quito. Biographical interviews with transnational families were held, thereby introducing dimensional, time, spatial and intergenerational factors into the analysis of transnational social protection strategies. This methodology also allows the "crossing" of family member narratives, shedding further light on how "interpersonal pacts" are formed and breakdown, as well as on intergenerational and gender relations (for further details see Oso & Suárez-Grimalt, 2017). Our selection of interviewees was based on gender and generation variables, as well as the type of transnational family, depending on who initiated the migratory process (mothers, fathers; children; or siblings).

The fieldwork also had a longitudinal dimension, as it was conducted over two time periods. Phase one of this research was carried out in 2008, coinciding with the outbreak of the financial crisis, although its impact was yet to make itself felt. It included biographical interviews with persons with family members in Spain, held in a district of southern Quito, together with interviews with key informants. This was followed by interviews with relatives of some of the people contacted in Quito, who had settled in Madrid. Eighteen people were interviewed during the initial phase of our fieldwork (11 women and 7 men).

In order to analyse the impact of the Great Crisis of 2008 on transnational social protection dynamics, a second phase of fieldwork took place in 2015. Contact was re-established with four families who had been interviewed in 2008. Some of the interviews were repeated in Quito and Madrid (eight in total), and new members of the same families were also interviewed for the first time. This longitudinal approach allowed us to monitor the biographical narratives of these families over time. The fieldwork was completed with interviews with other residents of the district, key informants and a number of returnees in Quito. In this second phase of the fieldwork 23 interviews (13 women and 10 men) were carried out.

A total of 41 interviews, with 33 people (8 persons were interviewed twice), were carried out over the course of the two fieldworks: 19 women and 14 men (15 of whom were members of the chosen four monitored families). A further phase of the fieldwork was initiated in 2021 in order to analyse the impact of the COVID-19 crisis. To date, one of the four monitored families has been interviewed (Oso, dir., 2007-2010; Oso, dir., 2015-2020; Oso, dir., 2021-2024).

We have illustrated the analysis of empirical data with the interviews of the monitored family cases, discussing those testimonies that best illustrate the articulation of interpersonal pacts and their infringements for TSP provision. The analysis is based on the "grounded theory" approach (Charmaz, 2005).

TRANSNATIONAL FAMILY	DESCRIPTION	FAMILY MEMBERS IN ECUADOR	FAMILY MEMBERS IN SPAIN	FAMILY MEMBERS INTERVIEWED; INTERVIEW PLACE AND DATE	
María	Family comprising the mother, Magdalena, her three daughters (María, single, Lucía and Ana, separated) and their three descendants	Mother, youngest daughter (Ana) and Ana's daughter	María, Lucía and their two daughters	Magdalena and younger sister, Quito, 2008/2015	María and Lucía Madrid, 2008/2015
Ana	Family comprising a divorced couple and two sisters (one is married and has two descendants)	Father, who is currently living with another woman (who has two children from a previous union) and has had a daughter with her	Ana, mother and sister, together with the husband and descendants	Ana's father, Quito 2008	Ana, Madrid, 2008/2015/2021
Manuel	Family comprising a divorced couple, four children and a grandchild	Mother, two middle sons and the youngest daughter	Manuel, eldest daughter and granddaughter	Mother, two middle sons and the youngest daughter Quito, 2008/2015	Manuel, Madrid, 2008/2015
Elvira	Family comprising a couple, three children (Elvira and another two) and two grandchildren	Father (in a union with another woman following the mother's death), elder brother (with his wife and son), youngest sister	Elvira, who lived in Madrid, but who was residing in Buenos Aires in 2017	Mother, eldest brother and Elvira's father, Quito, 2008-2015	Elvira interviewed in Madrid in 2008 and in Buenos Aires in 2017
TOTAL NUMBER OF MONITORED FAMILY MEMBERS	15				
OTHER PEOPLE INTERVIEWED	18				

Source: Authors' own.

Figure 1. Description of the transnational families interviewed

3. Migration, Domestic Care Work and the Welfare Paradox from a Transnational Approach

According to Eurostat, 49.1% of the Latin American migrant population resident in the European Union has settled in Spain (Bayona-i-Carrasco & Avila-Tàpies, 2019), where this group accounts for 3.1% of the total population and 40.2% of the foreign-born population (data from Spanish Municipal Population Census 2021, that is an administrative register that includes demographic data and, in the case of the migrant population, also includes those people with an irregular status. In Spain, registration is compulsory and it is a reliable source of foreign population data, as inclusion on this register is a requirement for access to healthcare and education). Spain's economic growth, migration policies drawn up in a context of post-colonial relations con visa regulations and less restrictive citizenship requirements for the Latin American population (hereafter LAC. For instance, the LAC population are entitled to Spanish citizenship after two years of legal residence in the country, compared to ten years for people from other regions such as Africa), together with a growing demand in Spain for domestic and care workers, have been the key triggers for the settlement of migrants, in particular women (Bertoli & Fernández-Huertas, 2013). In 2020, the feminisation rate stood at 56.9%.

Domestic service is the biggest sector of employment for female migrant workers, who account for 17.8% of the workforce. This figure is higher in the case of women from the LAC region, which provides 64.3% of the workers in domestic service (data from Spanish Labour Force Survey 2020, which include workers both in a regular and an irregular situation), in contrast to 24.9% of domestic and care workers from Europe, 3.8% from Asia, and 7.1% from Africa. In 2020, the number of domestic workers stood at 454,000 and it is estimated that around 63.7% of these employees are female migrants. The concentration capacity of the migrant population in this sector has been addressed in numerous studies which indicate that growth in this sector is linked mainly to household care requirements that are inherent to population ageing, changes in family structures and the increasing participation of women in the labour market (Moré, 2018). Scholars also agree that the spread of this activity is associated with the increasingly international nature of the gendered work division, which would explain why it includes a large number of female migrant workers (Barañano & Marchetti, 2016). The concentration of LAC workers is attributable to the fact that this group fits in with the image of an ideal postcolonial Spanish-speaking worker with catholic values that are perfectly aligned with domestic tasks in general and care work in particular (Castellani & Martín-Díaz, 2019).

The spread of public policies addressing long-term care based on cash transfers rather than social services is another key factor in the consolidation of the commodification of these tasks on care work to migrants (Picchi, 2016). These cash transfers act as a kind of subsidy given directly to the families, enabling them to acquire, in the private market, the social service deemed appropriate by the public system ("Economic benefit linked a social care service": 10.7% of long-term public care system users in 2019) or alternatively, close relatives providing care work ("Economic benefit for care in the family environment": 30.3% of users in 2019). A number of studies (Díaz & Martínez-Buján, 2018) have shown that the money granted for the care of family members through "Economic benefit for care in the family environment" is being used to pay private carers hired through domestic service, partly because there is no way of monitoring the way the money received is used. The co-pay mechanism linked to service-related financial subsidies has also contributed to this process. It was established in the light of the austerity policies introduced during the global financial crisis of 2008 and consists of users making a financial contribution that would complete the cost of the social service acquired through this programme. The issue is that co-payment is income-indexed and the principal limitation is that this amount is very high, even for those on average or low incomes.

As an example, the average cost of a place in a care home in Spain is 1,800 euros (IMSERSO, 2019). For a person with a high degree of dependence,

the maximum subsidy is 715 euros, provided that their monthly income does not exceed 565 euros. This user would have to pay the difference; in other words, 1,085 euros, more than their disposable income. The same is true of home-based care. The cost per hour for this service in an average sized municipality is 12.70 euros. A severely dependent person would receive a maximum of 70 hours per month (Martínez-Virto & Hermoso, 2021) at a cost of some 889 euros. Considering that according to Social Security data the average pension in Spain in 1,140 euros, co-payment for this service would be 30%, whereby the user would pay 267 euros for just two hours of care per day. As a result, many users are unable to pay the stipulated co-pay and resort to domestic service as a more economical and flexible way of filling their care requirements.

This situation indicates a contradiction between the philosophy underlying the design of public resources and the adverse effects that arise during their application. We have coined this process as a "welfare paradox" (in line with the terminology of earlier authors mentioned in the "Introduction") that impacts not only on the actual users, but also on the care strategies of the family environment and other provision areas. We posit that two clearly differentiated "welfare paradoxes" can be identified. On the one hand, the social resources referred to above are immersed in a "commodification paradox", in that they have proved incapable of decommodifying care, and are reduced to subsidies that complement the hiring of domestic service, or discourage the use of public social services due to the cost involved for beneficiary families. In other words, these measures fail to reduce individuals' dependence on the market. Back in the 1990s, feminist scholars were already calling for the need to highlight the "commodified" dimension in order to guarantee equal access to public resources (Orloff, 1993). If we include co-payment in the equation, stratification regarding who can or cannot access certain social services becomes even more pronounced. On the other hand, this contradiction interacts simultaneously with the strategies adopted by Spanish families in order to resolve their needs, and is also connected with the conception of a dualized welfare state that differentiates workers on the basis of their entitlement to labour and social rights. In this way, we find, at one extreme, with well-paid workers who have access to welfare protection and, at the other extreme, we have precarious, low-paid and unprotected workers. "This segmentation has configured a dualized welfare, which creates a class of "worker citizens", the "insiders", who contribute to the welfare state and are entitled to its benefit and a subclass of "working poor", the "outsiders", who have limited access to welfare benefits" (Castellani, 2020, p. 3). This segmentation is complemented with a further stratification layer among domestic migrant care workers comprising ethnicity, legal status and migration policies. All these processes create an "exclusion paradox" that blocks the migrant population's access to the social rights enjoyed by other citizens, as they experience greater difficulties in maintaining contributions throughout their labour trajectory.

This impact is particularly harsh in the case of domestic migrant care workers, who are also bound by a much more restrictive framework for the protection of their labour rights than other workers. This legal framework permits cease and desist dismissal (whereby the employing family can fire a worker at any time they deem fit); non-entitlement to unemployment benefit (even though they are legally employed); exclusion from the "Occupational Risks Prevention Law"; and the absence of work inspections in the private households that employ them (Molero-Marañón, 2020). This situation places them in a position of extreme vulnerability, particularly during periods of systemic crisis. During the Great Recession of 2008, even though many domestic care workers remained in work, the continuous wage cuts and worsening working conditions were clearly in evidence on multiple occasions (Hellgren & Serrano, 2017). During the COVID-19 crisis, when they suddenly became essential workers, due to the particular vulnerability to the virus of the elderly population, their working conditions were so precarious that they became the most vulnerable workers of the pandemic (ILO, 2020). The introduction of Spanish government's "extra subsidy for domestic service workers" for workers who had lost their jobs

or had seen their working hours reduced as a result of the health emergency, did little to ease the situation (the subsidy was up to 70% of their contribution-based earnings, up to a maximum of 950 euros, the minimum salary for 2020). The condition of being registered with the Social Security detracted from its universalising effect as it is estimated that 30% of domestic and care workers are in the underground economy (Díaz & Martínez-Buján, 2018). This situation together with the delay in its introduction (applications were not accepted until May 2020) and late payments (which in some cases extended to up to three months following application) have further worsened these workers' living conditions, forcing them to resort to informal means of support or seek aid from voluntary organisations in order to cover their essential needs (Díaz & Elizalde, 2021). Consequently, distanced from the "social pact" that formed the foundations for the welfare state and the inability of social protection mechanisms to act quickly and effectively in the light of an emergency, adopting transnational strategies for achieving social protection has become an essential resistance strategy.

The following section, based on our fieldwork, presents the arguments that show how these informal social protection strategies, aimed at offsetting the paradox of exclusion and the breach of the "welfare pact", are based on interpersonal agreements rooted in marriage/couple and intergenerational pacts and their infringements. The tension between relationships of solidarity and conflicts underlie the adoption of informal social protection strategies and their articulation with formal ones on a transnational scale, within the framework of the "welfare paradox".

4. Intergenerational and Marriage/Couple Pacts in the Articulation of Social Protection Strategies of Transnational Families

Transnational social protection strategies are based on intergenerational family pacts that may involve mothers/fathers, grandmothers/grandfathers, sons/daughters, brothers/sisters and nephews/nieces, underpinned by the sense of family solidarity and the obligation to provide welfare (Ayuso, 2012), as reflected in María's testimony:

> "For us, the mother and father are sacred. It's as if they were our children; like we have a lifelong obligation to repay them for everything they have given us. As they have protected us and continue to do so up till now, now that they are older and have worked so hard, it's time for the children to play their part" (María, Madrid, 2008).

Transnational protection strategies are also articulated through marriage/couple pacts. In addition to their legal basis, which implies certain welfare rights and obligations for the partners thereto, is also grounded in cultural values and gender imperatives. As Pateman explained (1988), the sexual contract is also the principle of the marriage contract. Indeed, both the intergenerational and marriage/couple pacts are conditioned by gender roles. As Cortés and Oso (2017) point out, women are expected to play a greater role in the provision of welfare through care, whilst men's contribution to the family welfare is essentially through their status as the "breadwinner". However, the roles in intergenerational pact are disrupted in the case of those families in which the woman acts as the pioneer in the migratory chain, leaving the children behind in Ecuador. The gender rules are broken when these women leave their role as "in-person carers" delegating it to a third person in the country of origin, and assume the responsibility for family welfare through remittances. This situation is occasionally sanctioned by the children left behind in the country of origin, who may perceive that the emotional dimension of the intergenerational pact has been broken. It occurs less in the case of men who migrate alone, leaving their wives and descendants behind in Ecuador, as they assume the traditional role of breadwinner.

Ana's mother was one of the women that pioneered the migration process following Ecuador's economic crisis at the end of the 20th century, travelling to Spain to work as a domestic and care worker. Her migration was part of an intergenerational TSP strategy in order to provide their daughters, who

initially remained in Quito with their father, with a university education, which is very expensive in Ecuador. The interview with Anna reveals our informant's perception of an emotional breakdown of the intergenerational pact after her mother migrated: "I was 13 when my mother emigrated, and my feelings were of abandonment" (Ana, Madrid, 2015). In turn, Manuel's children, who were left with their mother in Quito, see their father's emigration in a more positive light, even though he separated and settled down with a new family in Madrid (couple and daughter) and his wife accepted this new relationship. This is attributable to the fact that, even though he had broken the emotional dimension of the marriage pact with her, he upheld his responsibilities regarding the intergenerational welfare pact, sending remittances and therefore complying with his assigned role of breadwinner. Thanks to their father's financial support, all three children were able to attend private schools in Ecuador and undertake vocational training or university courses, as part of an education-based transnational social protection strategy. Nevertheless, the mother did not want to grant him a divorce, because she feared that, after the formal break-up of the marriage pact, he would marry the other woman. A new marriage formal pact could have an impact on the children's inheritance (intergenerational pact).

> "I'm not giving him the divorce, because he left to look after the family. He is with another woman, but she doesn't stop him from sending me money. If I want him to continue to support me financially, I can't say anything" (Manuel's wife, Quito, 2015)

It is therefore clear that emotions also play a part in intergenerational pacts. A connection can be drawn between the emotional ties and the strength of the pact: the severance of these ties, the loss of affection or family quarrels can articulate the agreements on which the transnational social protection is based.

Magdalena's oldest daughter (Lucía) was the first to leave the country, followed by her husband, their two daughters and her middle sister, María, the only one who remained single and had no children. Lucía separated from her husband in Spain, shortly after he arrived in Madrid. The two sisters and Lucía's two daughters lived together in an apartment and worked in the domestic, care and cleaning sectors. The situation helped Lucía, as a strategy of informal social protection, after the breakdown of her marriage pact and her ex-husband's lack of responsibility towards their daughters. The two sisters shared the care of the daughters, who were still young children at the time. María took care of her nieces, as part of an intergenerational family welfare pact. Likewise, at the beginning of the migration cycle, the two migrant sisters, Lucía and María, provided financial support for the family who had remained in Quito, acting as a buffer, as a measure of transnational informal social protection, in the face of the mother's ineligibility for a retirement pension and public healthcare; the medical expenses incurred by the premature birth of the niece; and the fact that both their sister Ana and her husband were unemployed in Quito. As a result, the remittances sent by the migrant sisters were the principal source of income for the household in Ecuador. In addition to the money they sent from Spain for the daily expenses, María sent money to build an apartment for herself at the top of the house where her mother lived. This apartment was occupied by her youngest sister's family (Ana, her husband and the niece) as a form of informal social protection in terms of the provision of housing.

> "My family depend on us 100 %. My mother doesn't work, nor does my sister. The youngest, my niece, was born prematurely and we paid for all the hospital fees from here. Now, her husband is unemployed after he was dismissed, and we have to feed the four of them and the child, who is still a baby. (…) We work for others to live (…) My mother no longer receives public healthcare, she now has a private doctor to treat her ailments" (Maria, Madrid, 2008)

However, the family in Ecuador is aware of the need to have a good relationship with Maria to ensure that she will not question the intergenerational pact that guarantees remittances:

> "Now that we are so far apart, my sister back in Ecuador is much more affectionate towards me, just like my mum. She is closer to us and respects us more. She used to say whatever she felt like, but now she bites her tongue. Firstly, because she says that we are the elder sisters, and secondly because they depend on us - because otherwise we would stop sending money. Because when I get mad, I stop everything" (María, Madrid, 2015).

The onset of the 2008 economic crisis entailed the restructuring of the social protection strategies of the transnational family. Both Lucía and María experienced a sharp reduction in their working hours in domestic service and cleaning, which led to a substantial decrease in their income. They were blocked by the "exclusion paradox". The two sisters had a row, which prompted María to move into a small apartment all by herself, in another area of Madrid where she was living in 2015. Lucia's financial situation was quite precarious, after the breakdown of the intergenerational pact with her sister. She had to turn to the church as a form of social protection in order to receive food. She also occasionally received remittances sent by her mother in Ecuador, as the economic situation of the family back in Quito had improved considerably (Magdalena was awarded a pension and her daughter found a job). María is still helping her nieces, giving them money for food and clothing whenever she sees that the financial situation is dire. The youngest niece had moved in with her in 2015, as a means of social protection, in view of the difficult situation that Lucía was going through. This indicates that intergenerational social welfare pacts are re-established over time, and that transnational social protection strategies are forged in accordance with the strength or breakdown of emotional ties, but also in the light of the impact of the "exclusion paradox".

The intergenerational support pact of descendants towards their parents, or uncles and aunts towards their nephews and nieces, should supposedly be stronger in the case of single family members. This is due to the conflict between intergenerational pacts and marriage/couple pacts, especially where descendants are involved. Some of our female informants explained how they had to conceal the remittances they sent to their parents and siblings from their partners. Indeed, parents' intergenerational pacts with their children take precedence over all others.

> "I send money to my mother every month without fail. My sister was off sick for around six months and she said 'I can't send money to mum, because those ten euros are for my daughters' milk'. Or don't have children, if you have a partner, as they won't let you send money back to the family. Because the money is for the family, in other word the husband, wife and their children. As I'm single, they expect more from me. She says: 'It's because you don't have any expenses; you're single; you don't have children'. It's like, you have to send money, because you just have to" (María, Madrid, 2008)

This clash between intergenerational and marriage/couple pacts is also reflected in the case of Elvira, who, at the start of her migratory experience sent back half her salary, working in the domestic/care and catering sectors, to her mother. The money was invested in expanding the family business in Quito as part of an intergenerational social protection strategy. This safeguarded her brother's job and led to an overall improvement in the family's circumstances, and also enabled her younger sister to study at university. However, following our informant's marriage to a fellow Ecuadorian she met in Madrid put a stop to these remittances over several years, and they were only renewed following the couple's separation:

> "At the beginning I would send money back to my mother, which enabled her to purchase computers to set up a business, and also to help my brother and my sister (...) There was a period when I didn't send anything —that was when I got married— because you have other responsibilities. We had a mortgage and he [in reference to her husband] and I had to find the money for our expenses (...) Later, when we separated, I was able to start sending money again (...)" (Elvira, Madrid, 2008)

Following the onset of the economic crisis in 2008, Elvira, faced with the "exclusion paradox", decided

to return to Ecuador in 2014. However, she is unable to find her place in Quito. After the death of her mother, her father embarks on a sentimental relationship with another woman (a new couple pact), which interferes with the intergenerational pact. Elvira does not accept this union and begins to claim her space, questioning the father's authority, leading to family rows. She decides to migrate again, this time to Buenos Aires. All this leads to the rupture of the intergenerational pact.

> "(…) for me it was not positive that she has built a bigger house, a better business (…) she no longer has the same respect that she had back then, the respect seems to have gone… (Elvira's father, Quito, 2015).

In turn, intergenerational pacts are a means of ensuring social protection when a marriage/partnership breaks down. This is the case of Lucía, who, after separating from her husband, turned to her sister María for help. To further illustrate this idea, we return to the family history of Ana, who initially stayed in Ecuador following her mother's migration. After the family regrouped, Ana's father felt unsettled in Spain and ended up going back to Ecuador. He and his wife separated and finally divorced, thereby resulting in the breakdown of the "marriage pact". In 2021, the father was living with a new partner, and the couple had a daughter three years ago. His partner has two older children from a previous relationship who are studying at university. Ana has made it clear to her father that he must cherish the relationship with his new daughter and her step-siblings, because they will have to care for him when he is older. The abuse that the three women of the family suffered (domestic violence) and his decision to settle permanently in Quito, brought about a breach of the marriage and the intergenerational contract between the father and the three women. Ana has urged her father to safeguard the new pact with the younger generation of his family and his partner in Ecuador as a means of social protection that will ensure he is cared for in the future: "I tell my dad: behave towards them as you didn't behave towards us – be a father. I tell him: take advantage of this opportunity, maybe they will look after you when you are old, because we won't be able to" (Ana, Madrid, 2021).

The breakdown of the marriage pact with the father lies at the heart of the precariousness the family experienced in the wake of the 2008 recession. The mother lost her job and had no form of social protection, because as a domestic service worker she was not entitled to any form of unemployment benefits, as we saw in the first part of the article. She was also blocked by the "exclusion paradox". Her husband, settled in Quito, was not sending money to Spain, instead employing his financial resources for his own upkeep in Ecuador. In this case, he fails to assume his role as breadwinner.

The intergenerational pact forged between the three women became the sole form of social protection open to the family in order to face the welfare paradox, providing various degrees of support. The elder sister, who got married in Spain, had two children and moved to another flat with her husband, acted as a permanent link in this chain of intergenerational social protection support. Thanks to her husband's financial support (a new marriage pact) in paying their own household expenses, the elder sister was able to finance the costs of the flat Ana and her mother lived in and also provided meals for her mother during the times of greatest hardship. Thanks to her sister's help, Ana was able to study at the university, and as a result of those studies our informant found a skilled job in Madrid that enabled her to send remittances to her mother. In turn, the mother cares for the grandchildren and helps with domestic chores, thereby supporting her elder daughter, whose working hours in the catering industry and the lack of state support for families in Spain make securing a life-work balance a challenging task.

Unlike the 2008 recession, which impacted most severely on the building industry, the COVID-19 crisis had a devastating impact on the catering sector, where Ana's elder sister and brother-in-law worked. Thanks to state social protection, and specific aid for workers who lost their jobs during the health emergency (the furlough scheme known as ERTE), the couple were able to get through the recession,

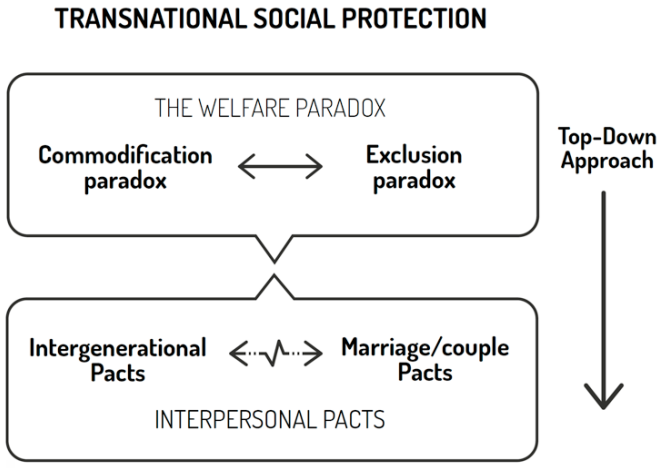
albeit on a far lower income. However, all members of the family who had settled in Spain, with the exception of Ana, caught coronavirus (her sister, brother-in-law, mother, niece and nephew). This situation placed Ana under great emotional stress, as she was in Madrid and extremely concerned about her family's health. Her mother also suffered, as she was afraid of dying due to her age and health problems. In turn, the elder sister suffered panic attacks because of concerns over the family's health problems and the after-effects of the disease she was personally experiencing. The fact that the family is separated, (Ana is alone in Madrid) and the lack of a solid social network, made it difficult for them to manage the health emergency:

> "We don't have a family network here (in reference to Spain). My sister is ill, and I'm not there to look after the children and take the medication to my mum. We don't have a strong, consolidated network. The fact that we don't have a family network, which I really miss, makes everything very difficult" (Ana, Madrid, 2021).

In turn, Ana was forced to go through the health emergency alone, working from home without seeing anyone, which also caused health issues. Faced with an emotional crisis, the intergenerational pact between women is activated, this time based on silence:

> "A close bond of solidarity has been forged between us: we look after one another, by not talking about what we're going through. How are you? Great, fantastic, even though we all have health problems... How are you? I'm fine. Sure? Yes Mum, I'm good" (Ana, Madrid, 2021).

This family's history highlights the way in which the pacts and their eventual breakdown shape informal social protection strategies, which are combined with others of a more formal nature. The breakdown of the marriage pact between husband and wife, together with the mother's low income and her non-entitlement to unemployment benefits as a domestic service worker ("exclusion paradox"), worsened the precarious position of the family in Spain after the 2008 crisis. In turn, an intergenerational social protection chain formed by the three women was forged to secure the younger daughter's education and offset the mother's unemployment and the elder daughter's difficulties in balancing work and family. In addition to this chain, a pact of silence was established after the COVID-19 crisis which guarantees emotional stability. In Ecuador, the crisis caused by COVID-19 is solved thanks to state social protection in the form of the father's pension, as well as the possibility of informal employment in the workshop, helped by his partner when she loses her job as domestic worker in Quito. In addition, the family members living in Ecuador occasionally call on Ana's sense of solidarity when they are in debt or need to make a particular purchase. Finally, due to the father's abusive behaviour, together with the distancing result from his decision to settle in Ecuador, and his failure to assume the role of breadwinner, any future care of the progenitor will depend on the strength of a new intergenerational pact with his youngest daughter and his new partner's other children in Quito.



Source: Authors' own.

Figure 2. The "welfare paradox" and the "interpersonal pacts" in transnational social protection

5. Conclusions

The first contribution of this article to the debate on TSP, is the application of the "welfare paradox" concept to show how formal social protection of long-

term care in Spain has resulted in the emergence of a labour market for domestic service which centres the activity of migrant women, in particular of Latin American origin ("commodification paradox"). Simultaneously, this process interacts with the difficulties these workers experience in accessing public social benefits, due to their non-contribution to the social security system and the legal framework that regulates domestic service ("exclusion paradox"). Expelling these workers from the "social welfare contract" in turn leads to the configuration of informal TSP strategies, particularly in contexts of crisis and social risk. A review of the public system for long-term care in Spain is also necessary in order to revert this situation. The elimination of the co-pay mechanism and the option of hiring professional carers through monetary transfers are two criteria that could be monitored in the mid-term. The ratification of ILO conventions 189 and 190 and the equation of the working conditions of domestic caregivers to other employees are key criteria in order to minimise processes of social exclusion.

A second contribution of this article is the analysis of the fieldwork conducted with Ecuadorian transnational families, which has revealed how these informal social protection strategies are supported by interpersonal pacts (intergenerational and marriage/couples) that are sustained by cultural norms and gender imperatives. These pacts are the result of relationships of solidarity, although they are not free from conflict. As a result, any breakdown in these pacts may also determine the nature of the transnational social protection strategies deployed. This tends to result in the activation of intergenerational pacts when marriage/couple pacts are weakened and vice versa. It is in the interplay of this tension between interpersonal pacts of solidarity, their breakdown and the exclusion in terms of the "social welfare pact" in the host country, that formal and informal social protection strategies come into play, applied within the transnational space.

Finally, the article highlights how the exchange of welfare between transnational families is linked to structural factors, which shape these practices. This top-down analysis has been overlooked in studies on TSP and is therefore our third contribution.

6. References

Amelina, A. (2017). *Transnationalizing inequalities in Europe: Sociocultural boundaries, assemblages, regimes of intersection*. London: Routledge.

Amelina, A. (2020). Theorizing large-scale societal relations through the conceptual lens of cross-border assemblages. *Current Sociology*, 69(3), 352-371. https://doi.org/10.1177/0011392120931145

Amelina, A., & Lutz, H. (2019). *Gender and migration: Transnational intersectional prospects*. London: Routledge.

Ayuso, L. (2012). El deber de apoyar a la familia. Una revisión del pacto intergeneracional de ayudas familiares en España. *Panorama Social*, 15: 143-158.

Baldassar, L. & L. Merla (2014). *Transnational Families, Migration and the Circulation of Care: Understanding Mobility and Absence in Family Life*. Abingdon: Rouledge.

Barañano, M., & S. Marchetti (2016). Perspectives on gender, migration and transnational work: joint work of social reproduction and care in Southern Europe. *Investigaciones feministas*, 7(1), 9-33. https://doi.org/10.5209/rev_INFE.2016.v7.n1.53094

Bayona-i-Carrasco, J., & Avila-Tàpies, R. (2019). Latin Americans and Caribbeans in Europe: A Cross-Country Analysis. International Migration, 58(1), 198-218. https://doi.org/10.1111/imig.12565

Bertoli, S., & Fernández-Huertas, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, 102, 79-100. https://doi.org/10.1016/j.jdeveco.2012.12.001

Castellani, S. (2020). On the fringes of Social Protection: New Southern European Labour Migration to Germany. *International Migration*. https://doi.org/10.1111/imig.12760

Castellani, S. & Martín-Díaz, E. (2019). Re-writing the domestic role: transnational migrants' households between informal and formal social protection in Ecuador and Spain. *Comparative Migration Studies*, 7(7). https://doi.org/10.1186/s40878-018-0108-0

Charmaz, K. (2005). "Charmaz grounded theory in the 21st century. Applications for Advancing Social Justice Studies". In N.K. Denzin, & Y.S. Lincoln (eds.), *The Sage Handbook of Qualitative Research* (pp. 507-536). Thousend Oaks: Sage Publication.

Cortés, A., & Oso, L. (2017). "Avecillas y pájaros en vuelo transnacional: Retorno, género y estrategias de movilidad e inmovilidad entre Ecuador y España", *Revista Española de Sociología* (RES), 26(3). Doi:10.22325/fes/res.2017.28

Daly, M. & J. Lewis (2000). The Concept of Social Care and the Analysis of Contemporary Welfare States. *British Journal of Sociology*, 51(2): 281-298. https://doi.org/10.1111/j.1468-4446.2000. 00281.x

Díaz, M., & Elizalde-San Miguel, B. (2021). La inevitabilidad de los empleos de cuidado: la crisis de la COVID como reflejo de las limitaciones sociales y jurídicas en el sector del empleo del hogar. *Migraciones* (forthcoming).

Díaz, M., & Martínez-Buján, R. (2018). Mujeres migrantes y trabajos de cuidados: transformaciones del sector doméstico en España. *Panorama Social*, 27: 105-118.

Hellgren, Z., & I. Serrano (2017). Transnationalism and Financial Crisis: The Hampered Migration Projects of Female Do-

mestic Workers in Spain. *Social Sciences*, 6(1), 8. http://dx.doi.org/10.3390/socsci6010008

Hochschild, A. R. (2000). Global care chains and emotional surplus value. In W. Hutton and A. Giddens, *On the edge: Living with global capitalism* (pp.130-46). London: Jonathan Cape.

ILO (2020, June 16). La COVID-19 pone en jaque el sustento de más de 55 millones de personas dedicadas al trabajo doméstico. *OIT Noticias*. https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_748117/lang--es/index.htm

IMSERSO (2019). *Informe Anual IMSERSO 2019*. Madrid: Vicepresidencia Segunda del Gobierno.

Korpi, W., & Palme, J. (1998). The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries. *American Sociological Review*, 63(5), 661-687. https://doi.org/10.2307/2657333

Kowalewska, H. (2021). Bringing Women of Board? Family Policies, Quotas and Gender Diversity in Top Jobs. *Work, Employment and Society*, 35(4): 735-752. https://doi.org/10.1177/0950017020971221

Lutz, H. (2018). Care migration: The connectivity between care chains, care circulation and transnational social inequality. *Current Sociology*, 66(4): 577-589. https://doi.org/10.1177%2F0011392118765213

Martínez-Buján, R. (2014). Domestic work count! Characteristics and transformations of the domestic Service in Spain. *Migraciones*, 36: 275-305. https://doi.org/10.14422/mig.i36.y2014.002

Martínez-Buján, R. (dir.) (2021-2024). *El modelo de cuidados de larga duración en transición: la articulación de programas comunitarios en el sistema público de bienestar tras la Covid-19*. Ministerio de Ciencia e Innovación, PID2020-114887RB-C33.

Martínez-Virto, L., & Hermoso, A. (2021). Hacia un modelo público de cuidados en la comunidad. *Revista Española de Sociología*, 30(2): a26. https://doi.org/10.22325/fes/res.2021.26

Molero-Marañón, M.L. (2020). Las trabajadoras del cuidado: por un futuro de trabajo decente. *Revista de Derecho Social*, 89: 33-64.

Moré, P. (2018). The Limits of Care. Organisation of the Household Assistance for the elderly in Madrid and Paris. *Revista Internacional de Sociología*, 76(1): e086. https://doi.org/10.3989/ris.2018.76.1.16.25

Orloff, A.S. (1993). Gender and the social rights of citizenship. *American Sociological Review*, 58(3): 303–328. https://doi.org/10.2307/2095903

Oso, L. (dir.) (2007-2010). *El impacto de la inmigración en el desarrollo: género y transnacionalismo*. Ministerio de Ciencia e Innovación, SEJ2007-63179.

Oso, L. (dir.) (2015-2020). *Género, movilidades cruzadas y dinámicas transnacionales, Ministerio de Economía y Competitividad*. Ministerio de Economía y Competitividad, FEM2015-67164-R.

Oso, L. (dir.) (2020-2024). *Care, Inequality and Wellbeing in Tansnational Families in Europe: A Comparative Intergenerational Study in Spain, France, Sweden and UK*. JPI-More Years, Better Lives, Ministerio de Ciencia e Innovación, PCI2021-121924

Parella, S. & Speroni, T. (2018). Las perspectivas transnacionales para el análisis de la protección social en contextos migratorios. *Autoctonía*, 2(1): 37-56. https://doi.org/10.23854/autoc.v2i1.59

Pateman, C. (1988). *The Sexual Contract*. Stanford: Stanford University Press.

Picchi, S. (2016). The elderly care and domestic services sector during the recent economic crisis. The case of Italy, Spain and France. *Revista de Investigaciones Feministas*, 7(1): 169-190. https://doi.org/10.5209/rev_INFE.2016.v7.n1.52067

Espainiako abiadura handiko trenaren ingurumen-balantze osoa

Espainiak munduko abiadura handiko trenbide-sare handienetan bigarrena du, Txinaren atzetik. Normalean, garraiobide jasangarria da, gasen emisioak eta energia-kontsumoa txikiak dituelako. Baina hori zenbakitan jartzeko saio gehienetan ez dira faktore guztiak kontuan hartzen. Kalkuluan azpiegitura eraiki eta mantentzearen karga ere kontuan hartzen badira, emaitzek oso zalantzan jartzen dute sarearen onura ekologikoa. Horrelako azpiegiturak planifikatzeko garaian, ezinbestekoa da ziklo osoa barne hartzea, erabakiak hartu aurretik.

Espainiako abiadura handiko trenaren (AVE) lehen linea 1992an jarri zen martxan, eta Madril eta Sevilla lotu zituen. Orduz geroztik, sarea asko handitu da, eta, gaur egun, ingurumen-profil konplexua du. Ingurumen-defizita linea berri bat eraikitzen denean hasten da, eraikuntzari eta mantentzelanei lotutako inpaktuengatik. Hala ere, martxan jartzen denean, ingurumen-kargak konpentsatu dezake, baldin eta bidaiariek gehiago kutsatzen duten beste garraio-mota batzuk utzi eta abiadura handiko trena erabiltzen badute. Neurketaren emaitza inpaktu garbi negatiboa izateak ingurumen-onura adierazten du.

Baina hasierako ingurumen-defizita konpentsatzeko behar den denbora aldatu egiten da korridorearen arabera. Adibidez, Kataluniako korridoreak —sareko garraioaren zati handi bat— 7 eta 12 urte bitartean konpentsatuko du inpaktua. Iparraldekoak eta beste korridore batzuek, aldiz, garraio-dentsitate txikia dutenez, zenbait inpaktu ez dira inoiz konpentsatuko azpiegituraren bizitza erabilgarriaren barruan.

Faktore batzuk faltan

Aztertutako faktoreen artean, ibilgailu pribatuen okupazio-tasak eta ibilgailu elektrikoetarako trantsizioa daude. Okupazio-tasa handia denean, txikiagoak dira abiadura handiko trenaren bidez garraiatzeak sortzen dituen onurak. Eta errepideko garraio pribatua nagusiki elektrikoa denean ere, trenaren abantailak murriztu egiten dira, elektrizitatea iturri berriztagarrietatik elikatzen bada. Garraioak etorkizunean egin dezakeen eboluzioaren arabera, zenbait eszenatoki aztertu dituzte ikertzaileek.

Espainiako AVEk inpaktua arinduko luke trafikoa modu kutsagarriago batzuetatik desbideratuko balu, baina ez ohiko trafikoa mantendu eta AVEn bidai-eskaera berriak sortuko balitu. Beraz, sarearen ingurumen-onura zalantzan jartzen da, bidaiari gutxi izaten baititu. Gaur egun, sareak ez du konpentsatzen errepide bidezko merkantzia-zirkulazioa. Abiadura handiko trenaren ingurumen-onurak nabariak dira hainbat urtez martxan egon ondoren, baina onurak murriztu egiten dira eskaera txikiko korridoreak gehitzen direnean.

«Abiadura handiko trenaren ingurumen-onurak murriztu egiten dira eskaera txikiko korridoreetan»

CO_2-emisioak murrizten badira ere, AVEk Espainian duen ingurumen-onura orokorra oso txikia da. Ordezko estrategiek ingurumen-onura handiagoak sor ditzakete, azpiegitura berriak eraikitzearen kostu eta eraginik gabe.

Munduko helburu klimatikoekin lerrokatzeko, arduradun politikoek AVE bezalako proiektu erraldoien ingurumen-ekarpen errealak ebaluatu behar dituzte. Halaber, beste estrategia eraginkor eta intentsibo batzuk hartu behar dituzte kontuan.

InpAkta 00 (2023) 245-267

https://doi.org/10.1016/j.retrec.2021.101035

Environmental balance of the high speed rail network in Spain: a life cycle assessment approach

Andoni Kortazar[1, 2], Gorka Bueno[2, 3], David Hoyos[2, 4]

[1] Department of Applied Economics V, University of the Basque Country UPV/EHU,; [2] EKOPOL, Research Group on Ecological Economics and Political Ecology; [3] Department of Electronics Engineering, University of the Basque Country UPV/EHU; [4] Department of Applied Economics III (Econometrics and Statistics), University of the Basque Country UPV/EHU

ABSTRACT: Spain has the world's second longest network of high speed rail lines built and in service. High-Speed Rail (HSR) is usually presented as a sustainable means of transport with huge potential to reduce greenhouse gas (GHG) emissions and energy consumption. The majority of studies carried out on this mode of transport have focused on analysing and estimating these savings in terms of network operation, but sometimes ignore the burdens associated with the construction of the infrastructure.

Based on the application of the Life Cycle Assessment (LCA) methodology, this work integrates into the analysis the construction and maintenance phases of the HSR lines in operation in Spain in 2016 together with their operation during that year, and verifies whether construction is justified in terms of reducing environmental impacts and energy consumption.

This article concludes that the construction of the Levante and Northern corridors is not justified in terms of energy savings and emission reductions due to the low demand and therefore the decision to build new HSR sections should be based on an analysis of demand so that only corridors with high transport demand are built. Furthermore, policymakers should consider other measures related to transport that would lead to considerable and rapid reductions in environmental impacts without the burden of building new infrastructures: e.g. reducing the demand for transport, increasing the occupation of private vehicles, promoting electric traction and the use of electricity from renewable sources.

1. Introduction

Spain has the world's second longest network of HSR lines, built and in service, surpassed only by China (European Court of Auditors, 2018). In relative terms, Spain is the world leader (Albalate and Bel, 2016). Popularly known as AVE (Alta Velocidad Española), Spanish High-Speed trains can travel at speeds of 250-300 km/h on international gauge lines of 1 435 mm, electrified at 25 kV and 50 Hz, on long distance routes (Ferropedia, 2016). Currently, these high-speed services are provided under dif-

* **Corresponding author / Harremanetan jartzeko:** Andoni Kortazar. Department of Applied Economics V, University of the Basque Country UPV/EHU, Spain. e-mail: andoni.cortazar@ehu.eus. **Contributed by** David Hoyos **editoreak egindako ekarpena**.
How to cite / Nola aipatu: Kortazar, Andoni *et al.* «Environmental balance of the High Speed Rail network in Spain: a Life Cycle Assessment approach», Research in Transportation Economics 90 (2021) 101035 (https://doi.org/10.1016/j.retrec.2021.101035)

ferent commercial brands besides the AVE, such as ALVIA, AVANT and AV-City. In 2005, the state company RENFE was divided into two companies, Renfe Operadora and Adif. The former is currently the only transport operator of passengers and merchandise in the Spanish railway sector, and is responsible for the maintenance and construction of trains. The latter, on the other hand, is the public company in charge of the exploitation of the railway infrastructure, and therefore it is in charge of the construction and maintenance of the HSR lines.

The first HSR line of the AVE between Madrid and Seville was inaugurated on 21 April 1992; since then, the network has spread throughout the country. Currently, four main corridors connect the capital city of Spain, Madrid, with other peripheral regions: Madrid-Catalonia, Madrid-León (Northern corridor), Madrid-Levante and Madrid-Andalusia. In total, 2 583 km[1] were already in service in 2016, and several more sections are currently under construction or projected. The majority of the AVE network is new construction with international gauge (1 435 mm), in order to solve the barrier effect historically created by the fact that conventional trains in Spain circulate on Iberian gauge (1 668 mm); variable-gauge trains that can travel on both networks are also used (Zembri and Libourel, 2017). There were twenty gauge changers in Spain in January 2017 (European Court of Auditors, 2018).

In general, investment in HSR can respond to several objectives, such as reducing the congestion of conventional networks, modernising the country, reducing oil dependency, improving interconnection with Europe, reducing travel times, or boosting territorial unity. These motivations directly affect the design of the network, its functionality and the effects caused on the environment (Albalate and Bel, 2016). The growing importance of environmental issues in public decision-making, especially in the transport sector, has forced the European Commission to insist on the need to apply a series of measures to limit the contribution of transport activity on climate change, calling for the strengthening of environmental assessments of any political initiative with major environmental impacts (European Commission, 2013).

In this context, HSR is usually presented as a sustainable means of transport with huge potential to achieve significant reductions of greenhouse gas (GHG) emissions and energy savings (California High-Speed Rail Authority, 2016; Jehanno *et al.*, 2011). However, the alleged savings have been questioned by some authors arguing that some of these environmental assessments focus on analysing and estimating these savings in terms of network operation, and fail to account for the burdens associated with the construction, maintenance and dismantling of the infrastructure (Bueno *et al.*, 2017; Chester and Horvath, 2010; Heather Jones *et al.*, 2016).

Given that any HSR project significantly alters the environment, a rigorous analysis of its environmental impacts and benefits becomes essential. In addition to the significant financial resources required, the construction of infrastructure megaprojects also require enormous amounts of natural resources such as concrete and steel, as well as terrain movements, resulting in a considerable consumption of energy and emission of pollutants into the atmosphere. Therefore, project evaluation requires a rigorous cost-benefit analysis to ensure social profitability (Flyvbjerg *et al.*, 2013) and, from an environmental perspective, to account for environmental impacts from cradle to grave, i.e. including the construction phase of the infrastructure and of the rolling stock (Baron *et al.*, 2011; Cour des Comptes, 2014).

The main objective of this article is to analyse the role of the Spanish HSR network in mitigating climate change and reducing energy consumption. Then, this work analyses the most significant environmental impacts and energy consumption associated with the construction, maintenance and operation of the Spanish HSR network under a 60-year lifetime horizon. This comprehensive analysis of the most important environmental burdens generated in the entire life cycle of the HSR network in Spain, are presented under the inventory of the following flows: CO_2eq (related to Global Warming); Cumulative Energy De-

[1] 3 402 km were already in operation at the end of 2019 (ADIF, 2020).

mand; PM_{10} (related to Human toxicity); SO_2 (Acidification, Human toxicity, Photochemical oxidation); NO_X (Acidification, Eutrophication, Human toxicity) and NMVOC (Ecotoxicity, Human toxicity, Photochemical oxidation, Ozone layer depletion)[2]. The geographic scope of the study is the entire network[3] in operation in 2016, which extends over 2 500 km along four main corridors.

After a brief introduction and description of the background of the project in section 1, section 2 presents the methodological basis for the study of the LCA of the HSR network in Spain. In section 3 the case study is presented: (1) the data and information necessary for the study is detailed (section 3.1); (2) the modelling approach for the environmental analysis is presented (section 3.2); (3) the characteristics and properties of the baseline scenario and sensitivity analysis are described (section 3.3). Section 4 and 5 presents and discusses the results, respectively. Finally, section 6 contains the conclusions drawn from this work.

2. Methodology

The analysis carried out in this work is based on the Life Cycle Assessment (LCA) methodology. This methodological tool is used to assess environmental impacts associated with all the stages of the life-cycle of a product or service, from raw material extraction to the processing of materials, infrastructure construction, use, maintenance and end-of-life treatments, based on a "cradle-to-grave" approach. LCA relies on the collection and analysis of all inputs (energy and materials consumption) and outputs (emissions, waste and by-products) of the system under study. The LCA of a product or service is often based on the use of exhaustive databases that compile the inventory of the life cycle of other products and processes already analysed.

Several studies have performed the LCA of HSR projects (Table 1). Chester and Horvath (2010) conducted a study on the California high speed line (CAHSR, 1 100 km), currently under construction, estimating that between six to eight years would be needed in order to balance the carbon footprint of the construction of the infrastructure, as long as the occupancy rate of the HSR remained higher than the rest of the modes of transportation. Chang and Kendall (2011) raised similar conclusions in their analysis of the San Francisco-Anaheim line (CAHSR, 725 km), considering that it would take six years to amortise GHG emissions, although more than 20 years might be needed if occupancy rates were lower than expected. Barnes (2014) also discussed the California line and showed that replacing some of the cement with fly ash in the production of concrete and the use of renewable energy for the HSR are the most feasible options for the CAHSR to be more efficient in the fight against climate change. Yue *et al.* (2015) performed the LCA on China's Beijing-Shanghai HSR, concluding that the operation phase of the HSR has a more significant contribution on the environment than the construction phase of its infrastructure, and that overall impacts can be substantially reduced if the consumption of coal to provide power is limited, the use of trains is optimised, fly ash is used, or if the construction of bridges and tunnels is limited. Banar and Özdemir (2015) performed the LCA of the Turkish HSR network determining that, of all the environmental burdens, 58% correspond to the construction phase and 42% to the operation phase.

Different works have carried out the LCA of HSR lines in Europe. Akerman (2011) studied the Swedish Europabanan line from an LCA perspective, determining that total emissions could be reduced by 0.55 million tonnes of CO_2 equivalent per year. Cornet *et al.* (2017) analysed the carbon footprint of the largest transport infrastructure project of the century in the United

[2] This analysis focuses on six main environmental indicators, namely GHG emissions, energy consumption, PM_{10}, SO_2, NO_X and NMVOC; however the reader should bear in mind that other environmental dimensions are also affected by the HSR construction and operation such as: habitat fragmentation, impacts on flora and fauna (affecting biodiversity), occupation of fertile land, landscape and visual impact, noise and vibrations, etc. In fact, impacts are generally similar along road and railways (Cour des Comptes, 2014; Dorsey *et al.*, (2015); Jehanno *et al.*, 2011). See Tables 1-9 and 14 of the Data in Brief.

[3] The section in Galicia is left out of this study because it does not meet the conditions of high speed when operating on Iberian gauge (Leboeuf M., 2018).

Kingdom, the HSR that will link London with different cities in the north (HS2), and argued that it will not contribute to a net reduction of CO_2 emissions.

Jones *et al.* (2016) analysed the total environmental impacts of the future Portuguese HSR line between Lisbon and Porto, finding that the operation of the HSR is the largest contributor to the total atmospheric emissions of the project (69% of CO_2, 76% of SO_2, 82% of PM_{10}), while the rest would correspond to burdens associated with the construction, maintenance and end-of-life treatment of the infrastructure.

In regards to the AVE network in Spain, to our knowledge no study has been conducted considering all the environmental burdens associated with the network's complete life-cycle. García Álvarez (2010) concluded that each passenger on the Madrid-Barcelona HSR line could contribute to an emissions reduction of 30 kg of CO_2, however failing to account for the infrastructure's construction burdens. Bueno *et al.* (2017), on the other hand, performed a limited environmental LCA of a new HSR line in the Basque Country, to be connected to the AVE network, that included the construction and maintenance phase of the infrastructure. They concluded that CO_2 emissions linked to infrastructure construction and maintenance would never be compensated, and that it would take 55 years of operation before net energy savings started.

Cuenot (2016) carried out for the UIC (International Union of Railways) a synthesis of the works on the main methodologies for the calculation of the carbon footprint of different railway lines, concluding that the most precise, transparent work, and with the best guarantees of offering the most reliable results, is that performed by Tuchschmid *et al.* (2011). This is the main reason why in this paper we follow the approach employed by Tuchschmid *et al.* (2011) for the calculations of the construction burdens of the Spanish HSR network. The environmental impact coefficients obtained in this study are collected in the Data in Brief.

Table 1
LCA studies about HSR projects

Project	Country	Reference	km	Included in the study
California HSR (CAHSR)	USA	Chester and Horvath (2010)	1 100	Infrastructure construction; operation
San Francisco-Anaheim (CAHSR)	USA	Chang and Kendall (2011)	725	Infrastructure construction
California HSR (CAHSR)	USA	Barnes E. (2014)	1 100	Design/Construction phase; Use phase
Beijing-Shanghai	China	Yue *et al.* (2015)	1 318	Vehicle manufacture, maintenance and disposal; infrastructure construction; operation
Turkish HSR	Turkey	Banar and Özdemir (2015)	888	Construction, maintenance and operation of railway infrastructure; production, maintenance and operation of train vehicles
Europabanan line	Sweden	Akerman J. (2011)	740	Construction, maintenance and operation of infrastructure; manufacturing and maintenance of vehicles
UK HS2	United Kingdom	Cornet *et al.* (2017)	556	Infrastructure construction; operation
Lisbon and Porto	Portugal	Jones *et al.* (2016)	297	Construction, maintenance, operation and disposal of railway infrastructure; production, maintenance, operation and disposal of train vehicles
Madrid-Barcelona	Spain	García Álvarez A. (2010)	621	Operation of the train
Y Basque	Spain	Bueno *et al.* (2017)	180	Infrastructure construction; operation

Source: Bellet Sanfeliu and Santos Ganges, 2016.

Figure 1. Spanish HSR network map (2016)

3. Case study

3.1. Data

Based on the application of the Life Cycle Assessment (LCA) methodology, this work integrates into the analysis the construction and maintenance phases of the HSR lines in operation in Spain in 2016 together with their operation during that year, and verifies whether construction is justified in terms of reducing environmental impacts and energy consumption.

The analysis of the environmental burdens associated with the construction of railway infrastructure requires, on the one hand, a precise description of the network (including a detailed inventory of all the materials used) and, on the other, information on the transport service provided by the infrastructure —measured in terms of passengers-km— and the modal shifts involved.

The functional unit for our analysis is the passenger transport service provided by the entire HSR network in one year of operation. For simplicity, the scenarios we examine assume that transport conditions remain unchanged during the infrastructure lifetime: annual passenger transport on the network, passengers shifted from other modes of transport, and the environmental impacts associated with both the operation of the HSR and the operation of other modes of transport (car, bus, conventional train and airplane). Freight transport is not included in this analysis, since the Spanish HSR network was designed exclusively for the transport of passengers (Bel, 2010).[4]

[4] It can be argued that the construction of a HSR network for passengers may release space for freight on the conventional network and that this effect should be taken into account. However, we find this unnecessary given the conventional network's current idle capacity.

Table 2
Details of the Spanish HSR network in 2016

Corridor	km	Tunnels (A)	%	Bridges & viaducts (B)	%	Total (A+B)	Total % (A+B)
Andalusia	646.8	46.5	7.2	46.5	7.2	93.0	14.4
Northern	445.2	46.7	10.5	20.9	4.7	67.6	15.2
Catalonia	883.0	95.3	10.8	66.3	7.5	161.6	18.3
Levante	607.9	59.2	9.7	36.9	6.1	96.1	15.8
Total	2 582.9	247.8	9.6	170.5	6.6	418.4	16.2

3.1.1. Layout of the AVE network

A detailed diagram of each branch of the Spanish HSR network is essential for the calculation of the environmental burdens of its construction and maintenance. This includes maps, routes, infrastructures, measurements, characteristics, etc. Most of the information publicly available consists of partial diagrams that do not allow an adequate characterisation. To remedy this, we conducted an exhaustive review of the routes through satellite and aerial images using Google Earth application, making it possible to characterise the infrastructure in detail[5] (Summary in Table 2).

The first line of the Spanish AVE network between Madrid and Seville was inaugurated on 21 April, 1992; since then, the network has spread throughout the country following a radial design around Madrid (Bel, 2010). Currently, there are four main corridors that connect different regions with the state capital. These four corridors are: Madrid-Catalonia, Madrid-León (Northern corridor), Madrid-Levante and Madrid-Andalusia. In total, 2 583 km were in service in the year 2016.

[5] The detailed description of the complete AVE network layout is available in the files attached to the Data in Brief document.

3.1.2. Passenger transport demand

An important problem that had to be addressed in this work was the absence of detailed quantitative data of HSR passengers and their corresponding travelling distances, essential to adequately assess the operation phase of the infrastructure.

Renfe Operadora, the only passenger operator in the Spanish railway sector, does not provide annual transport data –measured in pkm– on the Spanish HSR corridors, and only data for some specific connections are available (see Table 3) (Fundación de los ferrocarriles españoles, 2017). Every year, conventional long-distance passenger statistics are published without any reference to travelled distances, and with no specific reference to HSR passenger data. The only HSR passenger information available was provided by Galán *et al.* (2017), presenting data for passenger arrivals and departures for every HSR station in 2015. Similar information for the year 2016 was provided by García (2017). These sources, however, lack crucial information about the distances travelled. This information, however, allows the density of transport to be delimited throughout the infrastructure. For this purpose, a Python algorithm was programmed, which calculated a series of randomly generated matrices of passenger movements between stations compatible with the information provided by Galán *et al.* (2017) and García (2017), also considering the data provided for some specific connections by Fundación de los Ferrocarriles Españoles (2017).

Table 3
Passenger traffic in specific connections of the Spanish HSR network in 2016, provided by Fundación de los Ferrocarriles Españoles (2017)

Corridor	Connection	Annual passengers, published (in thousands)	Annual passengers, adjusted by Python algorithm (in thousands)
Andalusia	Madrid-Córdoba	880	884
	Madrid-Sevilla	2 545	2 556
	Madrid-Málaga	1 745	1 751
Northern corridor	Madrid-Valladolid	390	391
	Madrid-León, -Ourense, -Oviedo	939	940
Catalonia	Madrid-Zaragoza	1 373	1 384
	Madrid-Lleida	280	282
	Madrid-Tarragona	335	338
	Madrid-Barcelona	3 875	3 905
	Zaragoza-Barcelona	787	793
Levante	Madrid-Albacete	328	313
	Madrid-Alicante	1 394	1 330
	Madrid-Valencia	2 336	2 227

Table 4
Estimation of annual transport and density of transport for each corridor in 2016

Corridor	Length(km)	Passenger displacements (million)	Transport (million pkm, mean)	Transport (million pkm, standard deviation)	Length of average displacement (km)	Equivalent passengers over the complete infrastructure (million, mean)	Equivalent passengers over the complete infrastructure (million, standard deviation)
Andalusia	647	9.32	3 467	75.05	355	5.36	0.116
Northern	445	3.01	583	1.3	194	1.31	0.048
Catalonia	883	11.01	5 083	85.7	462	5.76	0.097
Levante	608	4.23	1 650	1.8	390	2.71	0.003
Total	2 583	27.57	10 783		391	4.17	

For any HSR line with n stations, characterised by its distance matrix (an $n \times n$ matrix with the distances between the n stations) and a vector with the annual travellers arriving/departing from each station, the algorithm allowed the random generation of an $n \times n$ matrix of movements among stations complying with the boundary conditions. Multiple executions (10^5) of the algorithm provided probabilities for passenger transport on the line under review that could be statistically treated.

Table 4 contains the estimation of transport for each of the four corridors of the Spanish HSR in 2016. For the Catalonia corridor, for example, while 11.01 million passenger journeys were recorded in 2016, 10^5 possible distributions of those movements compatible with the available record of passengers in stations provide a transport estimation with a mean value of 5 083 million passenger-km, and a standard deviation of 85.7 million passenger-km. This estimation is equivalent to a density of transport of 5.76 million passengers annually over the complete infrastructure, and implies that the average journey on the corridor is 462 km. For the whole Spanish network (2 583 km), the average journey is 391 km and density of transport is equivalent to 4.17 million passengers on the complete network. Figures 1-4 of de Data in Brief show the histograms for the transport calculation of 10^5 runs of the algorithm for each corridor.

3.2. Modelling approach

The net Environmental Impact (EI) balance of a new HSR infrastructure is provided by the comparison of the environmental impacts of all modes of transport in two alternative scenarios, one without the HSR, and the other with the HSR in service. Therefore, the net EI of constructing and operating a new HSR for a specific environmental parameter, such as CO_2eq or NO_x emissions, is mathematically represented as:

$$NetEI(\mathrm{t}) = EI_{withHSR} - EI_{withoutHSR} \quad (1)$$

Nevertheless, the calculation of the net EI does not require the explicit calculation of both transport scenarios, with and without the HSR line in operation, but just the differences between them: i.e. the impacts linked to the HSR transport of passengers in the scenario with HSR, and the impacts linked to the transport of passengers in other modes without the HSR that are shifted to the HSR when it is operative.

$$NetEI = EI_{\text{transportinHSR}} - EI_{\text{shifted transport from other modes to HSR}} \quad (2)$$

The modal shift of passengers from other more polluting modes of transport to the HSR is an environmental benefit of a HSR line. However, it needs to be balanced with the environmental impacts associated with its construction and to consider the induced transport. So, we calculate the net environmental impacts of any HSR line considering the phases of construction, maintenance and operation, and subtracting from it the environmental burdens linked to the transportation of passengers in other modes of transport that are shifted to the HSR. Mathematically, this can be expressed as:

$$NetEI = \sum EI^{HSR}_{Construction \wedge Maintenance} + \sum EI^{HSR}_{Operation} - \sum_i EI^{i}_{i \to HSR} \quad (3)$$

Where, *i* denotes the alternative passenger modes of transport to the HSR, i.e. aeroplane, conventional train, bus and private car. As the functional unit for the LCA is the transport service provided by the infrastructure of the HSR in a year of operation, in this calculation the environmental burdens associated with construction and maintenance of the infrastructure are evenly distributed along the considered lifetime horizon (60 years) in order to provide impact indicators in a yearly basis (Table 5).

According to equation [3], any HSR project will provide a net environmental benefit in a particular category if the annual environmental burdens associated to its construction, maintenance and operation are lower than the avoided burdens that transport shifted from other modes to the HSR would produce in a scenario without the HSR in operation (i.e. Net EI < 0). Consequently, net environmental benefits from any HSR line in a given category critically depend on the ability to attract substantial amounts of traffic from other modes of transport with higher emissions and energy intensity levels. In the following subsections, we examine the environmental burdens associated with the construction, maintenance and operation phases of the infrastructure (section 3.2.1) and the environmental savings produced by shifting passengers from other more polluting modes of transport to the HSR (section 3.2.2).

3.2.1. Environmental loads of the construction, maintenance and operation of the HSR

Environmental impacts linked to the construction, maintenance and operation of the HSR infrastructure were calculated as:

$$EI_{HSR} = \sum EI^{HSR}_{Construction \wedge Maintenance} + \sum EI^{HSR}_{Operation} \quad (4)$$

The Spanish HSR network had a total length of 2 583 km for the four corridors built and in operation in 2016. Out of the total railway network, 170 km (6.6%) correspond to viaducts and bridges, and 248 km (9.6%) correspond to tunnels, with the independent tunnels of Guadarrama standing out, with a length of 28 km each.

The environmental impacts associated with the construction and maintenance of the HSR network were calculated applying to each item of the infrastructure (bridge, tunnel, etc.) the corresponding impact coefficient following Tuchschmid *et al.* (2011). All the coefficients for each infrastructure's element and impact category are contained in the Data in Brief. As some infrastructure elements, such as rail for tracks and sleepers, have a 30 years lifespan, some construction burdens will occur after half the lifetime of the infrastructure has elapsed. Most maintenance burdens occur throughout the infrastructure lifespan, but our calculation considers them together with all construction burdens and assumes that they occur in the construction phase, for simplification. This deviation from real behaviour implies an overestimation of the years of operation needed for compensation (Tables 9 and 10). This error, however, is very limited, as total maintenance burdens are only 1% of construction burdens (Tuchschmid *et al.* 2011).

Table 5 provides the total environmental impact in each parameter linked to the construction and maintenance of the Spanish HSR network, subdivided by the corridors. These impacts are quantified in terms of annual tonnes or terajoules per kilometre of network.

Table 5
Annual environmental impact linked to construction and maintenance of the four corridors of the Spanish HSR network

	CO_2	CED	PM_{10}	SO_2	NO_X	NMVOC
	$t \cdot km^{-1} \cdot y^{-1}$	$TJ \cdot km^{-1} \cdot y^{-1}$	$t \cdot km^{-1} \cdot y^{-1}$	$t \cdot km^{-1} \cdot y^{-1}$	$t \cdot km^{-1} \cdot y^{-1}$	$t \cdot km^{-1} \cdot y^{-1}$
Andalusia corridor (647 km)	64.846	0.885	0.102	0.164	0.254	0.041
Northern corridor (445 km)	77.775	0.971	0.106	0.171	0.263	0.042
Catalonia corridor (883 km)	83.593	1.051	0.117	0.183	0.289	0.047
Levante corridor (608 km)	76.212	0.993	0.112	0.178	0.279	0.045

The environmental impacts are similar for all the corridors when measured in relative terms. Although the Catalonia corridor presents higher relative impacts and the Andalusia corridor lower relative impacts, deviations from the average for each corridor are below 15% in CO_2 emissions, 10% for CED, and below 7% for PM_{10}, NO_X and NMVOC emissions. The average for CO_2 emissions (75.85 $t \cdot km^{-1} \cdot y^{-1}$) is slightly higher than the range provided by Tuchschmid *et al.* (2011) for conventional networks in some countries (from 38.8 to 71 $t \cdot km^{-1} \cdot y^{-1}$), but slightly lower than the range provided by Baron *et al.* (2011) for some HSR networks (from 79 to 270 $t \cdot km^{-1} \cdot y^{-1}$).

The environmental loads associated with the operation of the HSR were obtained from the Ecoinvent database, version 3.7, processed with the open LCA

software (see Table 6). As the Ecoinvent database includes the environmental burdens linked to the construction and maintenance of the transport infrastructures and vehicles (roads and railways, cars, bus, conventional trains, aeroplanes, etc.), the construction and maintenance burdens of the HSR were previously removed from Ecoinvent results in order to avoid double counting.

Table 6
Impact category coefficients for modes of transport. Own elaboration based on Ecoinvent v3.7 (S1, S2, S3, S4, S5 refer to sensitivity analysis scenarios)

TRANSPORT MODE	Global Warming $gCO_2eq \cdot pkm^{-1}$	CED $MJ \cdot pkm^{-1}$	PM_{10} $g \cdot pkm^{-1}$	SO_2 $g \cdot pkm^{-1}$	NO_X $g \cdot pkm^{-1}$	NMVOC $g \cdot pkm^{-1}$
passenger aircraft, very short haul	159.37	2.41	0.04	0.22	0.72	0.10
passenger coach	49.44	0.82	0.03	0.05	0.47	0.05
passenger car mix (56% diesel, 44% petrol, 1,68 p/v)	187.39	2.86	0.11	0.31	0.46	0.17
passenger car mix (56% diesel, 44% petrol, 2,52 p/v) (S1)	124.92	1.91	0.07	0.21	0.31	0.11
passenger car mix (56% diesel, 44% petrol, 3,36 p/v) (S2)	93.69	1.43	0.05	0.16	0.23	0.08
electric passenger car, Spanish electricity mix, 1,68 p/v (S4)	92.74	1.99	0.12	0.36	0.29	0.07
electric passenger car, 100% renewable electricity mix, 3,36 p/v (S5)	28.15	0.70	0.05	0.09	0.09	0.03
passenger train, Spanish electricity mix	54.65	1.18	0.05	0.19	0.31	0.03
passenger train, 100% renewable electricity mix (S5)	27.12	0.73	0.04	0.05	0.22	0.03
passenger train, high-speed, Spanish electricity mix, (HSR infrastructure excluded)	32.91	0.86	0.03	0.15	0.10	0.01
passenger train, high-speed, 100% renewable electricity mix (HSR infrastructure excluded) (S5)	6.42	0.43	0.02	0.02	0.02	0.00

The same environmental parameters proposed by Tuchschmid *et al.* (2011) were obtained from the Ecoinvent v3.7: Cumulative Energy Demand (CED), as an indicator of primary energy consumption; Carbon dioxide equivalent emissions (CO_2eq), as an indicator of global warming; and Particulate matter (PM_{10}), Sulphur dioxide (SO_2), Non-methane volatile organic compounds (NMVOC) and Nitrogen oxides (NO_X) emissions as indicators of potentially severe consequences to human health and ecosystems.

3.2.2. Environmental loads associated with shifted transport to the HSR

Avoided environmental impacts linked to all the transport shifted from other modes to the HSR, were calculated as:

$$EI_{Avoided} = \Sigma_i EI^i_{i \rightarrow HSR} \quad (5)$$

The net environmental balance of any HSR infrastructure benefits from any environmental saving that may arise from its implementation, mainly due to the passengers shifted from other more polluting modes of transport. Betancor and Llobet (2015) estimated the origin of the passengers travelling through each of the HSR line sections, specifying percentages of shifted (and induced) traffic. This information is contained in Table 7.

Table 7
Shifted transport from other modes of transport to the HSR. Own elaboration based on Betancor and Llobet (2015)

	LINE SECTION					
	Madrid-Barcelona	Madrid-Zaragoza Zaragoza-Barcelona	Rest of Catalonia corridor	Andalusia corridor	Levante corridor	Northern corridor
From airplane	43.00%	2.67%	0%	45%	45%	0%
From bus	3.25%	1.33%	0%	2%	2%	5%
From car	16.07%	20.00%	45%	12%	15%	30%
From train	27.29%	49.33%	45%	26%	30%	35%
New demand	10.39%	26.67%	10%	15%	8%	30%

For any specific category, the avoided environmental impact linked to transport shifted from any mode in a section of the corridor to the HSR, is calculated by multiplying the corresponding transport density in the corridor (Table 4, measured in $pkm \cdot y^{-1}$) by the percentage for the shifted transport for that mode (Table 7) and by the corresponding impact category coefficient (Table 6; e.g. 159.37 $gCO_2eq \cdot p^{-1} \cdot km^{-1}$ for avoided CO_2eq emissions due to shifted air transport).

3.3. *Description of scenarios*

All the scenarios examined in this paper consider a fixed technological and socioeconomic context during the entire infrastructure's life cycle. This implies that important variables, such as transport demand, diverted traffic, vehicle occupation rates, energy intensities or electricity mix were considered invariable. The calculation period is established in 60 years, as it is the expected useful lifetime of most of the components of a railway network (Stripple and Uppenberg, 2010).

The Baseline Scenario that serves as a reference for this LCA departs from the following assumptions: 56% of the private vehicles in Spain are diesel cars, and 44% are petrol cars[6]; shifted transport from other modes to HSR follows Betancor and Llobet (2015); passenger transport in each corridor of the HSR infrastructure in every year of the lifetime according to section 3.2.2 (Table 4); electricity mix for Spain in 2017 (Ecoinvent v3.7) and, occupancy rate in private vehicles is 1.68 passengers per vehicle following the survey carried out by the S.G de Explotación in 2014 (Ministerio de Transportes, Movilidad y Agenda Urbana, 2018).

[6] Statistical data of the DGT (DGT, 2018) about the park of vehicles of Spain in 2018.

In order to respond to the uncertainty of some of the variables considered, five sensitivity scenarios were examined (S1, S2, S3, S4, S5). In each case, one of the variables assumed in the Baseline Scenario was modified in order to analyse its influence on the results. These scenarios do not have any influence on the construction and maintenance burdens, but they may have an impact on the burdens associated with the shifted transport from other modes and the operation of the HSR.

Table 8
Description of the scenarios

	Passenger car occupancy rate (persons per vehicle)	HSR demand	Passenger car	Electricity mix
Baseline Scenario	1.68	Baseline Scenario	56% diesel car, 44% petrol car	Spain 2017
S1	2.52	Baseline Scenario	56% diesel car, 44% petrol car	Spain 2017
S2	3.36	Baseline Scenario	56% diesel car, 44% petrol car	Spain 2017
S3	1.68	100% higher	56% diesel car, 44% petrol car	Spain 2017
S4	1.68	Baseline Scenario	Electric car	Spain 2017
S5	3.36	Baseline Scenario	Electric car	100% renewable

The first two scenarios analyse the influence of an increment of the vehicle occupancy rate. The S1 scenario assumes an occupancy rate of 2.52 passengers per vehicle, as estimated for Spain in 2020 by Adra *et al.*(2010). The second scenario (S2) analyses the influence of new mobility policies that institutions may promote in terms of vehicle sharing, specifically if the average occupancy rate of private vehicles were to double, 3.36 passengers per vehicle. The third scenario (S3) examines the influence of transport demand assuming that current demand doubles on all corridors. The fourth scenario (S4) explores the influence of transport electrification assuming that all private cars are electric. Finally, the fifth scenario (S5) considers an all-electric car scenario with an average occupancy rate of 3.36 passengers per vehicle, and cars, conventional trains and HSR powered with a 100% renewable electricity mix.

4. Results

This section presents the results obtained with the modelling carried out for the Baseline Scenario, as well as for each of the five alternative scenarios proposed for sensitivity analysis.

4.1. Baseline Scenario

Detailed information about the environmental balance of each corridor of the AVE network is contained in the Data in Brief. Table 9 summarises the main results in the Baseline Scenario for the Catalonia corridor, which supports 47% of total transport on the network (see Table 4). Table 10 summarises the main results in the Baseline Scenario for the whole Spanish HSR network.

A negative result in the net impact balance corresponds to a benefit in environmental terms. According to equation [3], a negative result implies that environmental burdens associated with the construction, maintenance and the operation of the HSR network are counterbalanced by the burdens avoided which are linked to the transport of those passengers that are shifted from other modes of transport to HSR.

Table 9
Environmental impact balance for the Catalonia corridor in the Baseline Scenario

Catalonia corridor	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
	$kt{\cdot}CO_2eq{\cdot}y^{-1}$	$TJ{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$
Infrastructure Construction & Maintenance	73.81	927.77	103.72	161.62	254.79	41.69
Operation (5.76 million passengers)	167.27	4370.85	172.39	779.85	524.49	32.17
Shifted transport in other modes	543.71	8900.62	288.55	1017.30	2101.56	393.85
Net Impact	-302.64	-3602.00	-12.44	-75.83	-1322.29	-319.99
Years required for compensation	11.76	12.29	53.57	40.84	9.69	6.92

Table 10
Environmental impact balance for the AVE network (four corridors) in the Baseline Scenario. (N.C: No Compensation in 100 years)

Spanish HSR Network	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
	$kt{\cdot}CO_2eq{\cdot}y^{-1}$	$TJ{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$
Infrastructure Construction & Maintenance	196.69	2 535.83	285.02	452.22	705.31	114.96
Operation (4.17 million passengers)	354.82	9271.65	365.68	1654.25	1112.56	68.24
Shifted transport in other modes	1161.63	18838.51	562.48	2088.87	4696.56	809.87
Net Impact	-610.13	-7031.03	88.22	17.60	-2878.69	-626.67
Years required for compensation	years	years	years	years	years	years
Four corridors (4.17 million passengers)	15	16	87	62	12	9
Andalusia corridor (5.36 million passengers)	9	11	96	57	8	7
Northern corridor (1.31 million passengers)	79	98	N.C	N.C	73	34
Catalonia corridor (5.76 million passengers)	12	12	54	41	10	7
Levante corridor (2.71 million passengers)	20	21	N.C	80	16	14

As most of the impacts linked to the construction and maintenance of the infrastructure occur before the line is put into operation, any new HSR network will start with an environmental deficit that will be compensated after a number of years of operation, if the operation impacts are less than the impacts avoided in other modes of transport. The number of years of operation needed to compensate that initial deficit will vary for each impact category. The exact moment in which these compensations begin is crucial information, especially regarding national objectives related to energy savings and emissions reduction deadlines. The years needed to provide such compensation in each impact category are provided in the last rows of Tables 9 and 10.

As can be observed in Table 10, the net environmental balance for the entire AVE network is negative in

all impact categories, except for PM_{10} and SO_2 emissions. CO_2eq, CED, NO_X and NMVOC impact categories need between 9 and 16 years to reach compensation, while SO_2 emissions need 62 years, and PM_{10} emissions (87 years) will not be compensated during the useful life of the infrastructure.

However, these global results vary significantly from corridor to corridor (see data in Tables 22-25 in the Data in Brief). All the corridors connect the periphery of the peninsula with Madrid in a radial design, but with very different conditions regarding transport density and the avoided transport mode mix. Then, it is essential to study the introduction of the HSR in a case-by-case analysis, taking a deeper look into the environmental performance of each of the corridors. By doing so, this work has detected elements that may go unnoticed in other more analytical and methodological studies (D'Alfonso *et al.*, 2015). This is the case of the burdens associated with the construction of the infrastructure. D'Alfonso *et al.* (2016) assume, based on other studies, that the construction of the HSR infrastructure adds an extra 5 g CO_2 per passenger-kilometre of transport served in the network. This value, however, depends absolutely on the amount of transport served by the network, which in the case of the Spanish HSR is much lower than that of other networks in the world (see Table 15). The data collected in Tables 4 and 5 allow for the calculation of the GHG footprint associated with the construction and maintenance of each of the corridors and for the whole network, taking into account the annual transport in each corridor (see Table 26 in the Data in Brief). The average footprint is 18.24 gCO_2eq/pkm for the whole network (3.6 times the value assumed by D'Alfonso *et al.* (2016)), but with large differences between corridors: 12.1 g in Andalusia, 14.5 g in Catalonia, 28.1 in Levante and 59.3 in the Northern corridor.

The Catalonia corridor (Madrid-Barcelona-France) supports the highest transport volume on the network and a high proportion of transport shifted from aeroplane, which provides similar results to the Andalusia corridor. CO_2eq, CED, NO_X and NMVOC impacts need between 7 and 12 years to be compensated; SO_2 needs 41 years, and PM_{10} is compensated in 54 years. Clearly, this corridor stands as the main contributor to the global warming emissions reduction of the entire Spanish network, as it is responsible for 44% of the total reduction of CO_2eq emissions. This corridor also stands as the main contributor to the reduction of energy consumption (51%) of the network.

The Andalusia corridor (Madrid-Sevilla-Málaga) requires between 7 and 11 years to compensate CO_2eq, CED, NO_X and NMVOC impacts. Around 57 years are needed to compensate SO_2, and no compensation (96 years) is achieved regarding PM_{10}, mainly due to the high proportion of transport shifted from aeroplane (45%). But the annual GHG emissions avoided by air transport (249 kt CO_2 annually) amount to 65% of the total emissions avoided, and are even much higher than the emissions linked to the transport induced in the HSR (15% of the total transport, 17.1 kt CO_2 annually).

Results worsen in the other two corridors, mainly due to the very low density of transport. In the Levante corridor (Madrid-Valencia-Alicante) CO_2eq, CED, NO_X and NMVOC impacts need between 14 and 21 years to be compensated; SO_2 (80 years) and, PM_{10} (143 years) emissions would not be compensated during the lifetime of the infrastructure.

The Northern corridor (Madrid-León-Zamora) offers the poorest results. A transport density of just 1.3 million passengers over the complete infrastructure gives rise to no compensation in the lifetime of the infrastructure in all of the impact categories analysed except for NMVOC, which would need 34 years.

4.2. *Sensitivity analysis*

A sensitivity analysis was carried out in order to check the influence of changes in certain variables over the total net environmental balance of the Spanish HSR infrastructure. Five alternative scenarios were analysed (see section 3.3 for a detailed description). The results are presented in Tables 11-14, in terms of years needed for compensation of environmental burdens linked to construction and maintenance of the infrastructure, for each of the four corridors in the AVE network.

Table 11
Results of the net balance in the Andalusia corridor for the sensitivity analysis. (N.C: No Compensation in 100 years)

	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
Years required for compensation (Baseline, years)	9	11	96	57	8	7
Years required for compensation (S1, years)	11	13	N.C	94	8	8
Years required for compensation (S2, years)	11	14	N.C	N.C	8	8
Years required for compensation (S3, years)	5	6	48	29	4	4
Years required for compensation (S4, years)	11	13	85	49	8	8
Years required for compensation (S5, years)	9	10	65	18	7	9

Table 12
Results of the net balance in the Northern corridor for the sensitivity analysis. (N.C: No Compensation in 100 years)

	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
Years required for compensation (Baseline, years)	79	98	N.C	N.C	73	34
Years required for compensation (S1, years)	N.C	N.C	N.C	N.C	N.C	48
Years required for compensation (S2, years)	N.C	N.C	N.C	N.C	N.C	61
Years required for compensation (S3, years)	40	49	N.C	N.C	36	17
Years required for compensation (S4, years)	N.C	N.C	N.C	N.C	N.C	67
Years required for compensation (S5, years)	N.C	N.C	N.C	N.C	N.C	N.C

Table 13
Results of the net balance in the Catalonia corridor for the sensitivity analysis. (N.C: No Compensation in 100 years)

	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
Years required for compensation (Baseline, years)	12	12	54	41	10	7
Years required for compensation (S1, years)	15	17	88	92	11	9
Years required for compensation (S2, years)	17	20	N.C	N.C	12	10
Years required for compensation (S3, years)	6	6	27	20	5	3
Years required for compensation (S4, years)	17	16	47	33	11	10
Years required for compensation (S5, years)	17	17	64	24	11	12

Table 14
Results of the net balance in the Levante corridor for the sensitivity analysis. (N.C: No Compensation in 100 years)

	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
Years required for compensation (Baseline, years)	20	21	N.C	80	16	14
Years required for compensation (S1, years)	22	25	N.C	N.C	17	16
Years required for compensation (S2, years)	24	27	N.C	N.C	17	17
Years required for compensation (S3, years)	10	11	72	40	8	7
Years required for compensation (S4, years)	24	25	N.C	71	17	17
Years required for compensation (S5, years)	21	22	N.C	36	16	19

It can be seen that the results are somewhat sensitive to the occupancy rate of private vehicles. Impact compensation requires a few more years when the occupancy rate of private vehicles is higher (S1 and S2). Average occupancy rate could increase during the lifetime of the infrastructure, since there are currently several institutional initiatives aimed at promoting vehicle sharing in society, such VAO road lanes exclusively for high occupancy vehicles in Madrid. If that rate is doubled, 3.36 people per vehicle (S2), keeping the number of HSR travellers constant, every environmental indicator worsens with respect to the Baseline Scenario. This is due to the fact that the greater the occupation of private vehicles, the lower the environmental impact per person and kilometre of this mode of private transport. Thus, under S1 and S2 circumstances, attracting passengers to the HSR from private vehicles provides less beneficial effects on the net environmental impact of the HSR project.

Section 3.1.2 described the estimation of the HSR transport demand for each corridor in the year 2016 for the Baseline Scenario. The network considered in this study includes all the HSR corridors in operation at the end of 2016. Under the assumption that passenger transport demand may increase in the future, the S3 scenario considers that the annual transport demand in each corridor doubles, with the rest of the parameters remaining equal. From an environmental point of view, a rise in demand has two opposite impacts: a positive impact, when it relates to a modal shift from more polluting modes of transport, such as air or road; and a negative impact, when it is due to induced (new) demand.

As shown in Tables 11 to 14, doubling passenger demand improves the environmental performance of all the corridors, reducing the compensation period by more than half, given that the previously mentioned positive effect is greater than the negative effect. However, it is important to denote that these results are highly sensitive to the magnitude of the induced demand: (1) induced demand should not be higher than 70% of the new demand in the corridor of Andalusia, in order to achieve an improvement in CO_2eq emissions; (2) in the Catalonia corridor all new demand could be induced demand; (3) it should not exceed 63% in the case of the Levante corridor; and (4) it should stay below 44% and 34% in the Northern corridor in order to obtain an improvement in terms of CO_2eq and CED, respectively.

The S4 scenario explores a situation in which private vehicles are electric. It is reasonable to consider a progressive penetration of electric motion in the automotive sector, which will reduce future environmental impacts in road transport. The influence of this variable on the annual results of the AVE is noteworthy. In this scenario the burdens associated with the construction, maintenance and operation

of the Spanish HSR four corridors do not vary with respect to the Baseline Scenario, but the loads associated with shifted transport are now reduced. All the indicators worsen their balance with respect to the Baseline Scenario, except for PM_{10} and SO_2 that improve, although the former is almost residual. When electrifying road transport, the environmental burdens significantly reduce in this mode, so the benefits from diverting traffic from road to rail are lower. These environmental benefits are even lower if electric vehicles are powered with electricity from renewable sources. This context is explored in the fifth and last scenario (S5), where an occupancy rate of 3.36 passengers per vehicle is considered, together with the assumption that all private vehicles are powered with electricity from renewable sources. Compensation years for this scenario are similar to those provided by the previous one, S4, as a worsening derived from doubling vehicle occupancy tends to compensate with the benefit derived from the fact that the HSR would also operate on 100% renewable electricity.

5. Discussion

The main factor behind the net environmental balance of the Spanish HSR network is clearly the density of total demand as long as it is capable of diverting traffic from more polluting modes of transport (air or road) rather than inducing new demand. Vehicle occupancy rates, electrification of road transport and electricity mix have also been found to have a significant impact on the environmental balance of the network.

In line with the findings of many economic analyses (e.g. Albalate and Bel, 2011; Betancor and Llobet, 2015; De Rus, 2011), the performance of the Spanish HSR network is clearly hampered by its low passenger demand. In 2016 the Spanish HSR network transported the equivalent of just 4.17 million passengers over the complete infrastructure; significantly lower than the transport density supported by other networks in the world, as can be observed in Table 15.

Table 15
Passenger transport density in national HSR networks in 2016. Own elaboration based on UIC Railway Statistics (2016)

		Network length, maximum speed between 160 and 250 km/h	Network length, maximum speed greater than 250 km/h	Transport	Average density of transport (passengers over complete infrastructure)
National network	Year	(km)	(km)	(Mpkm)	(Mp)
Spanish AVE	2016	669	2 503	15 059	4.75
France	2015	0	2 043	49 980	24.46
Germany	2016	1 511	994	27 213	10.86
Italy	2012	2 767	653	12 794	3.74
China	2017	12 276	20 305	577 635	17.73
Japan (CJRC)	2017	0	553	54 756	99.02
Japan (EJR)	2017	0	1 194	23 371	19.57
Japan (WJRC)	2017	0	813	21 023	25.87
Taiwan (THSRC)	2016	0	350	10 488	29.97

The Spanish HSR network is also hampered by the fact that it allows only for passenger transport, which prevents attracting potentially more polluting traffic from road freight transport. As argued by Akerman (2011, p. 208), "HSR investments may not be justified for the passenger markets alone."

The LCA of Spanish HSR is quite robust in showing that the launch of the Catalonia and Andalusia corridors of the Spanish HSR network has led to a net environmental benefit in CO_2eq after nine-twelve years of operation. However, it also shows that results worsen as the network expands to corridors with lower demand (Levante or Northern corridors). This also coincides with transport economics literature (see e.g. Albalate and Bel, 2011).

However, two questions also need to be addressed in order to evaluate the environmental performance in absolute terms: firstly, is there a significant annual reduction in CO_2eq emissions and energy consumption provided by the Spanish HSR operation? And secondly, how does this reduction compare with other alternative strategies for managing transport passenger demand?

Overall environmental impact reductions derived from the HSR operation in Spain are very limited, if not negligible. Spain's transport sector had a total volume of direct emissions of 85.9 Mt CO_2eq in 2016 (European Environment Agency, 2016). The HSR network presents in the Baseline Scenario an annual net balance of -610.13 kt CO_2eq (Table 10), or less than 1% of emissions linked to transport. Regarding energy consumption, the net balance of -7 031.03 TJ (0.17 million toe) in cumulative energy demand is less than 0.5% of the energy consumed by the transport sector in Spain in 2016 (International Energy Agency, 2017). In other words, the Spanish HSR network's capacity to mitigate climate change and reduce oil dependency is clearly insufficient in the current context of the global environmental crisis, which requires a drastic reduction in GHG emissions. In cost-efficiency terms, it is important to consider that the total investment in AVE over the last 25 years has exceeded 50 000 million euros.

Table 16

Comparison of the environmental impact balances for the Catalonia corridor in the Baseline Scenario and a motorway corridor of the same length with scenarios of occupancy doubling (HW1), vehicle electrification (HW2), and occupancy doubling and renewable electricity vehicles (HW3)

	Global Warming	CED	PM_{10}	SO_2	NO_X	NMVOC
Net Impact balance	$kt.CO_2eq{\cdot}y^{-1}$	$TJ{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$	$t{\cdot}y^{-1}$
Catalonia corridor (Baseline Scenario)	−303	−3 602	−12	−76	−1 322	−320
HW1 - toll highway of same length, 883 km (6 million $v{\cdot}y^{-1}$, double occupancy)	−858	−12 924	−601	−1 586	−2 132	−752
HW2 - toll highway of same length, 883 km (6 million $v{\cdot}y^{-1}$, electric vehicles, renewable elec.)	−1 218	−13 336	−177	−426	−2 709	−871
HW3 - toll highway of same length, 883 km (6 $Mv{\cdot}y^{-1}$, double occupancy, ren. elect. vehicles)	−1 431	−19 350	−684	−1 718	−4 287	−1 114

From a sustainable mobility perspective, it is important to bear in mind the existence of other alternatives that could further reduce environmental impacts in the transport sector without the need to build new infrastructures (Hoyos, 2009). To illustrate this, Table 16 collects the environmental impact balance of the Catalonia corridor of the AVE network in comparison with three other scenarios applied to a

motorway corridor of the same length (883 km). According to traffic statistics from the Ministry of Public Works and Transport (Ministerio de Fomento, 2017), the average daily flow on motorway toll roads in Spain (2 550 km) was 16 471 vehicles (heavy vehicles excluded) in 2017, which is equivalent to an annual transport density of 6.01 million light vehicles annually. While the Catalonia corridor gives rise annually to a reduction of emissions of 303 kt CO_2eq, the doubling of light vehicle occupancy on a motorway toll road with the same length (883 km) and under average Spanish traffic conditions would provide an annual reduction of 858 kt CO_2eq (HW1 in Table 16). If those light vehicles were electrified, the reduction would increase to 1 218 kt CO_2eq (HW2); and up to 1 431 kt CO_2eq if vehicles were also powered with electricity from renewable sources, and with double the average occupancy (HW3).

6. Conclusions

Climate change and oil scarcity have received increasing attention in transport policy. In this context, HSR has often been presented as a sustainable mode of transport, having a leading role in the European Commission's environmental goal of net-zero GHG emissions by 2050, due to its potential contribution to energy savings and GHG emissions reductions. In this paper, it is assessed the environmental performance of the Spanish HSR network by means of LCA under 2016 traffic conditions. Results show that the construction loads of the Spanish network are not disproportionate, as they remain within the lower limit of the range of construction burdens found with other HSR lines. Although these construction loads are not excessive, the net environmental balance of the entire network in the Baseline Scenario, without being detrimental in almost all indicators, is modest: an annual emission reduction of 610 kt CO_2eq, of 7 031 TJ of CED, of 2 879 t NO_X and of 627 t NMVOC, together with an annual increase of 88 t $PM_{10,}$ of 17 t SO_2. This modest balance means that the infrastructure requires a minimum number of years of operation to offset the initial loads associated with the construction: between 9 and 16 years in all the environmental categories studied except for SO_2 (62 years) and PM_{10} with 87 years. In absolute terms, it means a reduction in CO_2eq emissions equivalent to less than 1% of the annual transport emissions in Spain in the base year (2016), together with a reduction in primary energy demand which is less than the equivalent of 0.5% of annual energy consumption in the transport sector.

The environmental balance varies according to the network corridor considered. The corridors of Catalonia (5.76 Mp) and Andalusia (5.36 Mp) present a slightly better balance than the total average, managing to compensate the initial construction loads in less than 7-12 years in all the analysed categories except for PM_{10} and SO_2. The Northern corridor (1.31 Mp) would not be able to compensate the initial loads in the whole time of operation, and the Levante corridor (2.71 Mp) would need around 14-21 years (except for SO_2 with 80 years and PM_{10} which would not be compensated). According to these results, the construction of the Levante and Northern corridors is not justified in terms of energy savings and emission reductions. Thus, in line with the findings on cost-benefit analysis, the decision to build new HSR sections should be based on the analysis of demand in order to build only those sections that ensure a high demand, that is, to build only those corridors that connect centres with high demographic density (De Rus, 2011).

The sensitivity analysis confirms that the main factor that conditions the net environmental balance is the density of the transport. The density of transport served by the network in 2016, measured in terms of equivalent passengers over the entire network layout (4.17 million passengers) is much lower than the transport served by the French network (24 Mp), the Japanese networks (between 20 and 99 Mp), China (18 Mp) and Taiwan (30 Mp) (Table 4 and 15). Initially, every HSR project starts operation with an environmental deficit that can be compensated only if transport demand is sufficiently high and it comes, sufficiently, from other modes of transport, minimising new induced demand.

Doubling the quantity of passengers, which is not very likely in the medium term, shows that the

results improve in all the corridors, even in the Northern corridor, and the amortisation of several indicators will be given within the term of the infrastructure's useful life. But the nature of this increase in transport demand is a relevant aspect that conditions the results in a crucial way. It is essential that this increase in total demand for HSR is not new induced demand in its entirety; otherwise, the adverse consequences on the environment will increase. In other words, the environmental performance of HSR improves if traffic is diverted from more polluting modes of transport and induced transport remains low. According to results, induced demand should not exceed 44-70% of total demand (except for Catalonia corridor) if positive effects are to be found in terms of CO_2eq emissions. Similar results are found regarding other pollutants and energy consumption. So, from a sustainability perspective, our findings suggest that public institutions should focus on increasing the current levels of passenger demand in the AVE network, prioritising the diversion of existing demand from planes and private cars.

Finally, the United Nations has recently declared a climate emergency under the latest scientific evidence on the consequences of climate change, emphasising the urgent need to be carbon neutral in 2050 and to achieve a 45% reduction in emissions by 2030 in order to keep the rise of temperature below 1.5 degrees Celsius by the end of the century (UNFCCC, 2019). In this context, policymakers should also consider other measures related to transport that, in application of the transport hierarchy (Hoyos *et al.*, 2016), would provide considerable and rapid reductions in environmental impacts without the burden of building new infrastructures: e.g. reducing the demand for transport, increasing the occupation of private vehicles, promoting electric traction and the use of electricity from renewable sources.

7. Acknowledgements

The authors acknowledge the funding provided by the University of the Basque Country UPV/EHU under research grant US15/11, by the Department of Education of the Basque Government through Grants IT-1359-19 (UPV/EHU Econometrics Research Group) and IT-1365-19 (UPV/EHU Ekopol Research Group), and by the Spanish Ministry of Economy, Industry and Competitiveness through grant ECO2017-82111-R.

8. References

ADIF, 2020. Adif - Líneas de alta velocidad [WWW Document]. ADIF Alta Velocidad. URL http://www.adifaltavelocidad.es/es_ES/infraestructuras/lineas_de_alta_velocidad/lineas_de_alta_velocidad.shtml (accessed 2.14.20).

Adra, N., Michaux, J.-L., André, M., 2010. Analysis of the load factor and the empty running rate for road transport. Artemis - assessment and reliability of transport emission models and inventory systems [WWW Document]. ResearchGate. URL https://www.researchgate.net/publication/277183200_Analysis_of_the_load_factor_and_the_empty_running_rate_for_road_transport_Artemis_-_assessment_and_reliability_of_transport_emission_models_and_inventory_systems (accessed 2.10.20).

Akerman, J., 2011. The role of high-speed rail in mitigating climate change – The Swedish case Europabanan from a life cycle perspective. Transportation Research Part D: Transport and Environment 16, 208–217. https://doi.org/10.1016/j.trd.2010.12.004

Albalate, D., Bel, G., 2016. Evaluating High-Speed Rail: Interdisciplinary Perspectives. Taylor & Francis.

Albalate, D., Bel, G., 2011. Cuando la economía no importa: auge y esplendor de la alta velocidad en España. Revista de Economía Aplicada 19.

Banar, M., Özdemir, A., 2015. An evaluation of railway passenger transport in Turkey using life cycle assessment and life cycle cost methods. Transportation Research Part D: Transport and Environment 41, 88–105. https://doi.org/10.1016/j.trd.2015.09.017

Barnes, E., 2014. California High Speed Resilience to Climate Change, in: Center for Earth Systems Engineering and Management, Course Project Report Series.

Baron, T., Martinetti, G., Pépion, D., 2011. Carbon footprint of high speed rail. International Union of Railways (UIC), Paris.

Bel, G., 2010. España, capital Paris. Destino.

Bellet Sanfeliu, C., Santos Ganges, L., 2016. The high-speed rail project as an urban redevelopment tool. The cases of Zaragoza and Valladolid. Belgeo. Revue belge de géographie. https://doi.org/10.4000/belgeo.18153

Betancor, O., Llobet, G., 2015. Contabilidad financiera y social de la alta velocidad en España. FEDEA: Área de infraestructuras y Transportes.

Bueno, G., Hoyos, D., Capellán-Pérez, I., 2017. Evaluating the environmental performance of the high speed rail project in the Basque Country, Spain. Research in Transportation Economics 62, 44–56. https://doi.org/10.1016/j.retrec.2017.02.004

California High-Speed Rail Authority, 2016. California High-Speed Rail Sustainability Report 54.

Chang, B., Kendall, A., 2011. Life cycle greenhouse gas assessment of infrastructure construction for California's high-speed rail system. Transportation Research Part D: Transport and Environment 16, 429–434. https://doi.org/10.1016/j.trd.2011.04.004

Chester, M., Horvath, A., 2010. Life-cycle assessment of high-speed rail: the case of California. Environ. Res. Lett. 5, 014003. https://doi.org/10.1088/1748-9326/5/1/014003

Cornet, Y., Dudley, G., Banister, D., 2017. High Speed Rail: Implications for carbon emissions and biodiversity. Case Studies on Transport Policy. https://doi.org/10.1016/j.cstp.2017.08.007

Cour des Comptes, 2014. La grande vitesse ferroviaire: un modele porte au-dela de sa pertinence. Paris.

Cuenot, F., 2016. Carbon Footprint of Railway Infraestructure: comparing existing methodologies on typical corridors. UIC-ETF 51.

D'Alfonso, T., Jiang, C., Bracaglia, V., 2016. Air transport and high-speed rail competition: Environmental implications and mitigation strategies. Transportation Research Part A: Policy and Practice 92, 261–276. https://doi.org/10.1016/j.tra.2016.06.009

D'Alfonso, T., Jiang, C., Bracaglia, V., 2015. Would competition between air transport and high-speed rail benefit environment and social welfare? Transportation Research Part B: Methodological 74. https://doi.org/10.1016/j.trb.2015.01.007

De Rus, G., 2011. The BCA of HSR: Should the government invest in high speed rail infrastructure? Journal of Benefit-Cost Analysis 2, 1–28.

DGT, 2018. DGT- Parque de vehículos en España [WWW Document]. DGT- Tablas estadísticas. URL http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/parque-vehiculos/tablas-estadisticas/ (accessed 2.14.20).

Dorsey, B., Olsson, M., Rew, L.J., 2015. Ecological Effects of Railways on Wildlife, in: Ree, R. van der, Smith, D.J., Grilo, C. (Eds.), Handbook of Road Ecology. John Wiley & Sons, Ltd, pp. 219–227.

European Commission, 2013. EU energy, transport and GHG emissions - trends to 2050: reference scenario 2013. European Commission, Luxembourg.

European Court of Auditors, 2018. A European high-speed rail network: not a reality but an ineffective patchwork [WWW Document]. URL https://www.eca.europa.eu/Lists/ECADocuments/SR18_19/SR_HIGH_SPEED_RAIL_EN.pdf (accessed 11.15.18).

European Environment Agency, 2016. GHG Proxy Inventory-SPAIN 2016 [WWW Document]. European Environment Agency. URL http://cdr.eionet.europa.eu/es/eu/mmr/art08_proxy/envwwefiw (accessed 4.24.19).

Ferropedia, 2016. Líneas de Alta Velocidad (LAV) en España - Ferropedia [WWW Document]. URL http://www.ferropedia.es/mediawiki/index.php/L%C3%ADneas_de_Alta_Velocidad_en_Espa%C3%B1a (accessed 4.25.19).

Flyvbjerg, B., Bruzelius, N., Rothengatter, W., 2013. Megaprojects and Risk: An Anatomy of Ambition, First Edition. ed. Cambridge University Press, United Kingdom.

Fundación de los ferrocarriles españoles, 2017. Atlas of High Speed Rail in Spain.

Galán, J., Alameda, D., Abad, J.M., 2017. AVE: Estas son las estaciones de alta velocidad más y menos utilizadas de España | Economía | EL PAÍS [WWW Document]. URL https://elpais.com/economia/2017/03/01/actualidad/1488362770_011434.html (accessed 3.20.18).

García Álvarez, A., 2010. Energy Consumption and Emissions of High-Speed Trains. Transportation Research Record: Journal of the Transportation Research Board 2159, 27–35. https://doi.org/10.3141/2159-04

García, P., 2017. Las ocho estaciones del AVE en España con menos de 150 pasajeros al día [WWW Document]. El Independiente. URL https://www.elindependiente.com/economia/2017/11/12/estaciones-ave-espana-menos-viajeros/ (accessed 6.9.18).

Hoyos, D., 2009. Towards an operational concept of sustainable mobility. International Journal of Sustainable Development and Planning 4, 158-173. https://doi.org/10.2495/SDP-V4-N2-158-173

Hoyos, D., Bueno, G., Capellán-Pérez, I., 2016. Environmental assessment of high-speed rail. Evaluating High-Speed Rail: Interdisciplinary Perspectives 7, 119.

International Energy Agency, 2017. Statistics | Spain - Total Primary Energy Supply (TPES) by source (chart) [WWW Document]. URL https://www.iea.org/statistics/?country=SPAIN&year=2016&category=Key%20indicators&indicator=TPESbySource&mode=chart&dataTable=BALANCES (accessed 4.23.19).

Jehanno, A., Palmer, D., James, C., 2011. High Speed Rail and Sustainability.

Jones, H., Moura, F., Domingos, T., 2016. Life cycle assessment of high-speed rail: a case study in Portugal. The International Journal of Life Cycle Assessment 22. https://doi.org/10.1007/s11367-016-1177-7

Leboeuf, M., 2018. High Speed Rail: fast track to sustainable mobility. International Union of Railways (UIC).

Ministerio de Fomento, 2017. Boletín estadístico online - Información estadística - Ministerio de Fomento [WWW Document]. URL https://apps.fomento.gob.es/BoletinOnline/?nivel=2&orden=06000000 (accessed 4.23.19).

Ministerio de Transportes, Movilidad y Agenda Urbana, 2018. Anuario estadístico 2018. Capítulo 8. Tráfico.

Stripple, H., Uppenberg, S., 2010. Life cycle assessment of railways and rail transports - Application in environmental product declarations (EPDs) for the Bothnia Line. Swedish Environmental Research Institute, Sweden.

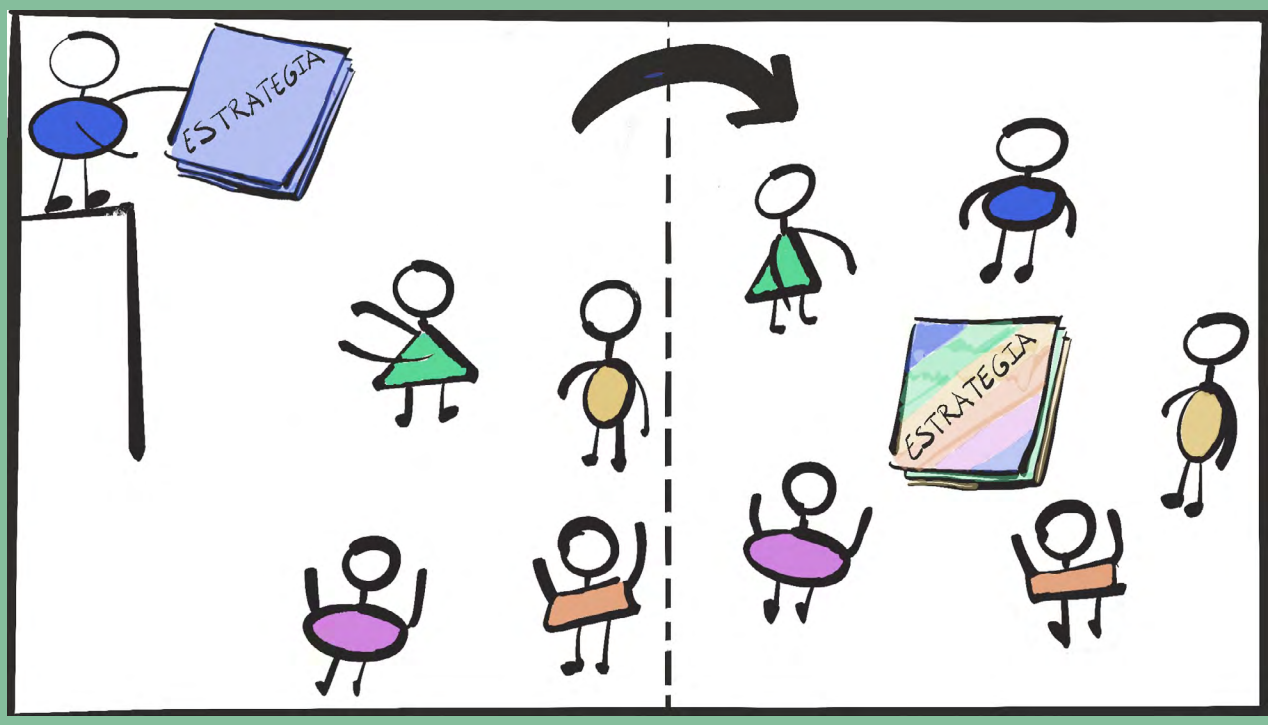
Tuchschmid, M., Knörr, W., Schacht, A., Mottschall, M., Schmied, M., 2011. Carbon Footprint and environmental impact of Railway Infrastructure.

UIC, 2016. RAILISA STAT UIC [WWW Document]. URL https://uic-stats.uic.org/ (accessed 2.1.20).

UNFCCC, 2019. UN Chief Calls on G7 Leaders to Tackle Climate Emergency | UNFCCC [WWW Document]. https://unfccc.int/news. URL https://unfccc.int/news/un-chief-calls-on-g7-leaders-to-tackle-climate-emergency (accessed 8.28.19).

Yue, Y., Wang, T., Liang, S., Yang, J., Hou, P., Qu, S., Zhou, J., Jia, X., Wang, H., Xu, M., 2015. Life cycle assessment of High Speed Rail in China. Transportation Research Part D: Transport and Environment 41, 367–376. https://doi.org/10.1016/j.trd.2015.10.005

Zembri, P., Libourel, E., 2017. Towards oversized high-speed rail systems? Some lessons from France and Spain. Transportation Research Procedia, World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016 25, 368–385. https://doi.org/10.1016/j.trpro.2017.05.414



Miren Larrea artikulu egilearen ilustrazioa

RIS3: Europako erregioetan espezializazio adimenduneko estrategiak garatzeko, ezinbestekoa da erregioz azpiko gobernuak ere inplikatzea

Erregioetako industria biziberritzeko, gero eta ohikoagoa da tokiaren indarguneak eta ikerketa-lerro estrategikoak identifikatzeko prozesu partizipatiboak abiatzea. Prozesuetan botila-lepo errepikakorrak izaten dira, ordea, eta haiek ebazteko zenbait gako proposatu dituzte ikertzaileek: aldundiak, garapen agentziak eta udalak inplikatzea, gatazkak bideratzeko baliabideak eskuratzea eta giza zientzietako ikertzaileak integratzea.

Europako Batzordeak, eredu ekonomiko berri bat lortze aldera, herrialde eta erregio guztiei eskatu die identifikatu ditzatela berrikuntzan eta industrian beste herrialdeen aldean dituzten indarguneak eta abantailak. Eta, horretan oinarrituta, zehaztu dezatela zein izango diren lehentasunez finantzatu beharreko berrikuntza-lerro estrategikoak. Espezializazio adimenduneko estrategia horiek aplikatzeak globalki lehiakorragoak egingo ditu erregioak.

AEBren aldean Europak ekoizpen industrialean duen arrakalarekin kezkatuta ezarri ditu Batzordeak espezializazio adimenduneko estrategiak (RIS3), baina ez da erraza praktikan gauzatzea, parte hartzeko prozesu konplexuak izaten baitira. Ikertzaileek proposatzen dute erregioz azpiko gobernuek, RIS3an eskumenik ez duten arren, funtsezko zeregina izan dezaketela, harreman zuzena izaten baitute prozesuan inplikatutako eragile askorekin. Gertutasun eta harreman-sare horiek baliatzeak gobernantza erraztu dezake.

Arrakastarako hiru gako

Ikerketaren arabera, prozesua eraginkorra izan dadin, ezinbestekoa da aurretik eragileen arteko sareak sortzea, elkarlanean aritu daitezen politika horiek formulatzeko garaian. Horrek erregioko gobernuen eta erregioz azpiko gobernuen arteko komunikazio zuzena eta eraginkorra eskatzen du.. EAEko lau adibide aztertu dituzte ikertzaileek maila anitzeko gobernantzaren arrakastarako gakoak identifikatzeko: Eusko Jaurlaritzak bideratutako prozesu bat, foru-aldundi batek bideratutakoa, eskualde mailakoa eta udal mailakoa. Horiek denak aztertuta, prozesuan azaldu diren arazoak gainditzeko gakoak eman dituzte.

Bestetik, enpresetako eragileek, ikertzaileek eta eragile sozialek gaitasun handia izan dezakete eskualdearen berrikuntza-indarguneak identifikatzeko. Beraz, gomendatzen dute administrazioak goitik behera ezarritako erabakiak ekiditea, eta behetik gorako ikuspegia erabiltzea. Alabaina, ikuspegi horrekin jokatzeko, ezinbestekoa ikusten dute botere-dinamikak alboratuko dituen gobernantza-eredu bat ezartzea. Prozesua konplexua izaten da, eta gatazkak sor daitezke erabakiak hartzeko unean. Beraz, ikusi dute funtsezkoa dela prozesua abiatu aurretik gatazkak ebazteko baliabideak bilatzea, prozesua arrakastatsua izango bada.

Azkenik, ikerketak agerian utzi du giza zientzietako ikertzaileek prozesuan parte hartzea ere gakoa izan daitekeela: beren ezagutza transmititu eta prozesua elkarrizketara bideratzeko jarrera hartu dezakete, eta gobernantza-mekanismo berritzaileak bultzatu. Aldaketarako eragile aktiboak izanik, murriztu eginen dute politiken eta inplementazio praktikoaren arteko arrakala.

InpAkta 00 (2023) 268-284

http://dx.doi.org/10.1080/09654313.2016.1174670

Overcoming policy making problems in smart specialization strategies: engaging subregional governments

Miren Estensoro and Miren Larrea

Orkestra-Basque Institute of Competitiveness and University of Deusto

ABSTRACT: Since the concept of Smart Specialization was launched, an effort has been made to clarify and establish criteria for its implementation. Part of the difficulties in implementing Research and Innovation Smart Specialization Strategies (RIS3) is their emphasis on bottom-up approaches, which are required because there are public and private stakeholders that are better positioned than governments to find the domains in which the region is likely to excel. Regions must shift towards a new generation of industrial policy and the difficulties for advancing in this direction are already visible. Designing and implementing a smart specialisation strategy at regional level: Some open questions. The centrality of entrepreneurial discovery in building and implementing a smart specialisation strategy. Efforts to implement smart specialization in practice – leading unlike horses to the water. The paper is based on four cases related to governance and learning for smart specialization in the Basque Country (Spain) and presents three main lessons learnt. The first has to do with connections between regional and sub-regional governments in order to construct networks of territorial actors that can act as the senses of governments in the territory. The second is about the challenge of handling complexity and conflict and the third is about the integration of social researchers in RIS3 processes.

1. Introduction

The concept of Smart Specialization and the related policy known as Research and Inno- vation Smart Specialization Strategies (RIS3) have become critical in European regional innovation and development policy since Dominique Foray (one of the founders of the concept) and a group of experts known as the "Knowledge for Growth" expert group was set up a few years ago to advise the European Commission (Capello, 2014).

The smart specialization argument originally emerged out of the discussion of the transatlantic produc-

* **Corresponding author / Harremanetan jartzeko:** Miren Estensoro. Orkestra-Basque Institute of Competitiveness and University of Deusto, Donostia-San Sebastián, Spain. e-mail: mestenso@orkestra.deusto.es. **Contributed by** Miren Estensoro **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Estensoro, Miren *et al.* «Overcoming policy making problems in smart specialization strategies: engaging subregional governments», *European Planning Studies*, 2016 (http://dx.doi.org/10.1080/09654313.2016.1174670)

tivity gap (Foray, David, & Hall, 2009). The focus here was initially on the role played by transatlantic differences in R&D intensity in explaining growth differentials. As McCann and Ortega-Argilés (2011, 2014) point out, the first explanation, called the "structural effect", emphasizes transatlantic differences in industrial structure. In particular, the EU's industrial structure is disproportionally characterized by traditional, middle and low-tech sectors, and this implies a lower capacity to translate R&D into productivity gains. The second explanation, known as the "intrinsic effect", argues that even within the same sectors, European firms exhibit a lower ability to translate R&D into productivity gains or other types of investment (McCann & Ortega-Argilés, 2014). According to these authors, the limited knowledge-transmission mechanisms between sectors and firms and also between regions, caused by the heterogeneity of the EU integration process, hinder the ability to spread the benefits of new R&D-related technologies across all sectors and industries.

Since the concept of Smart Specialization was launched and this debate began, an effort has been made to clarify and establish criteria for taking it into practice. It is currently playing a central role in the development of a reformed European Cohesion Policy and the Europe 2020 agenda, which is based on the principles of "smart growth", "green growth" and "inclusive growth" (Capello, 2014; Foray, 2014; Iacobucci, 2014; McCann & Ortega-Argilés, 2011, 2014). Landabaso (2014) underlines how, since 2009, and in the aftermath of the global financial and economic crisis, the concept of smart specialization has echoed in several European and global economic institutions, such as the Organization for Economic Cooperation and Development (OECD). We present this perspective in a nutshell by quoting Landabaso (2014, p. 132):

> Smart specialisation implies that a member state or region identifies and selects —on the basis of a bottom-up and top-down priority setting process— a limited number of priorities for knowledge-based investments focusing on regions' strengths and comparative advantages. This approach will hopefully help regions realise their innovation potential and refocus their industrial and knowledge assets in the direction of emerging industries and services and international markets.

Since the concept was launched, academics have aimed to help regions to start these pro- cesses. But especially in 2014 and 2015, when some results from these experiences became available, there has been a shift in the literature towards understanding the difficulties faced by regions in their implementation processes and how such difficulties can be over- come. In this article we interpret implementation as the process of putting into practical effect in the policy making process the concepts and frameworks elaborated for smart specialization and RIS3. It includes stages prior to the approval of the strategy and also later stages.

Following Coffano and Foray (2014), most regions are moving from the "easy-to-do" structural analysis to "hard-to-do" entrepreneurial discovery and they are struggling to achieve this. Authors such as Kroll (2015) and Ortega-Argilés (2012) have already addressed this challenge and our aim is to contribute to this by sharing lessons learnt in the Basque Country.

The reflections on the Basque case help to argue that some of the difficulties faced by governments in the implementation process are derived from the affirmation that RIS3 is neither "coffee for all", nor "picking winners from above" (Boschma, 2014; Capello, 2014; Foray, 2013; Foray *et al.*, 2009; Foray, David, & Hall, 2011; Landabaso, 2014; McCann & Ortega-Argilés, 2014). It is not about selecting firms or sectors, but about the research and innovation activities and the generic technology(ies) that can help a regional economy to diversify into higher value-added markets (McCann & Ortega-Argilés, 2011, 2014). Two of the main difficulties when implementing RIS3 are their bottom-up approaches and the aim to integrate private and public stakeholders (Foray *et al.*, 2009, 2012).

The implementation of these RIS3 requires regions to shift towards a new generation of "industrial policy" and the difficulties related to those challenges

are already visible (Coffano & Foray, 2014; Iacobucci, 2014; Kroll, 2015). In order to respond to the challenges of RIS3, we will argue that governments need to develop a network of territorial actors that can act as the senses of governments in the territory. This has to do with the model of governance in the region and we will also argue for the integration of sub- regional governments in these networks, which is consistent with Barca (2009) when he proposes a place-based approach where the responsibility for policy design and implemen- tation is allocated among different levels of government supported by both contractual relations and trust, with a role being played by special-purpose institutions such as agencies and public–private partnerships. The interpretation of RIS3 from a process per- spective and the role of social researchers as active agents of change in this field will also be proposed.

2. Problems with implementation

In the following section, the main implementation problems referred to in the recent lit- erature on RIS3 are emphasized in order to set a framework for the lessons learned from the case of the Basque Country. In this literature, there is a pattern that focuses on bottom- up approaches as a source of difficulties and on learning processes and the construction of capabilities, on the one hand, and the construction of new modes of governance, on the other, as ways to overcome such difficulties. In the following paragraphs, we examine pre- vious studies that have led us to such an interpretation.

2.1. *Difficulties with the implementation of bottom-up strategies*

Entrepreneurial discovery distinguishes smart specialization from traditional industrial and innovation policies (Landabaso, 2014). These discovery processes are a "main source of information about the new activities of exploration and transformation that are likely to be prioritized" (Coffano & Foray, 2014, p. 35).

Entrepreneurial discovery processes require governments to play a different role from the one played in traditional industrial and innovation policies, which includes facilitating bottom-up processes together with more traditional top-down procedures (Coffano & Foray, 2014; Foray *et al.*, 2009; OECD, 2012). From this perspective, prioritization decisions are expected to be driven by entrepreneurial knowledge and decisions, not by policy-makers' dreams (Coffano & Foray, 2014). This is an interactive and dynamic process in which market forces and the private sector discover and produce information about new activities and the government assesses the outcomes and empowers the actors that are most capable of realizing the potential (Landabaso, 2014).

Smart specialization was not conceived as a strategy for imposing specialization by means of top-down government planning. Rather, it was seen as being driven by a process of discovery and learning on the part of entrepreneurs, who are the best positioned agents to search for the right types of knowledge (McCann & Ortega-Argilés, 2011, 2014). Several authors have underlined the difficulties in the development of these processes.

McCann and Ortega-Argilés (2014) underline how, first of all, such a shift requires transparency in order to ensure public accountability. Coffano and Foray (2014) state how a policy design based on entrepreneurial discovery requires new models of incentives in order to encourage firms to elicit information and bring their own knowledge to the policy-makers, to induce entrepreneurs to come forward with their knowledge.

Camagni, Capello, and Lenzi (2014) and Capello (2014) argue that a bottom-up process of strategy definition runs the risk of a possible misallocation of public resources. The difficulties with prioritizing when defining this strategy are also mentioned by Capello (2014).

Iacobucci (2014) accepts that entrepreneurs are in a better position than government officials to identify opportunities and argues that bottom-up processes of entrepreneurial discovery will inevitably result in the proliferation of promising domains.

Still, without conscious moderation and guidance, a bottom-up approach seems to conflict with the idea of identifying a regional "strategy" that, in his opinion, at least at the beginning, must rely on a top-down approach (2014). Even when the stakeholders are invited to par- ticipate in the process, a top-down approach will be present when deciding which stake- holders will be involved and in the final decision on the chosen specialization domains, which depends on the regional government that is leading the strategy.

This position that argues for, but still considers the difficulties of, bottom-up approaches is also taken by McCann and Ortega-Argilés (2011). They refer to the need for smart specialization strategies to engage with local elites in order to extract local knowledge and to tailor the policy. But they also mention information asymmetries and principal–agent problems associated with engagement with local elites, together with externalities. Boschma (2014) argues for the need to prevent local elites from assuming monopolistic positions, which calls for a flexible but continuously monitored policy implementation process.

Table 1 synthetizes the main contributions considered in the previous paragraphs on this challenge.

Table 1
Implementation of bottom-up strategies

Type of challenge	Problem for implementation	Authors
Implementation of bottom-up strategies	Lock-in with respect to local historical specialization	Capello (2014)
	Monopoly positions of local stakeholders/elites lock-in and corruption as potential threats	Kroll (2015), Boschma (2014), Iacobucci (2014)
	Proliferation of promising domains and difficulties with prioritizing	Iacobucci (2014), Capello (2014)
	Misallocation of public resources whenever local interests and local political needs may set unfeasible industrial targets and risky innovation strategies	Camagni *et al.* (2014), Capello (2014)
	Lack of co-investment between public and private initiatives	McCann and Ortega-Argilés (2014)
	Top-down government planning; "blind giant"	McCann and Ortega-Argilés (2011, 2013, 2014), Foray and Goenaga (2013)

Source: authors' elaboration.

2.2. *The role of learning and capability building*

Some of the main arguments on how to overcome the challenge posed by bottom-up approaches relate to the development of capabilities as a goal and learning processes as a means for achieving that.

Coffano and Foray (2014) argue that the implementation of smart specialization pol- icies "requires good institutions and strong policy capabilities at the regional level". Some authors claim that one of the difficulties with learning is that there have been incon- sistencies in the academic production on RIS3 that have led to confusion when regional authorities have tried to implement the concept in practice (Capello, 2014; Kroll, 2015). Kroll (2015, p. 3) states that "early RIS3 policy guidelines provided little in the way of helping regional policy-makers

to make sense of local complexity and dynamism to the extent needed for solid policy-making". In order to deal with the learning challenge, the European Commission has established an interregional learning platform where regional and national governments exchange policy practices (Landabaso, 2014; McCann & Ortega-Argilés, 2014). Moreover, the EU Commission has produced a number of meth- odological guides and materials, as well as several tools and techniques (innovation vou- chers, assistance with patenting activities, etc.). It has also mobilized, for the first time, more than a hundred top academics and policy advisers to help regional and national governments with the production of their RIS3 in order to bridge the dangerously widen- ing gap in the European Union between practitioners and academics.

Foray and Goenaga (2013) emphasize the experimental nature of the policy process and conclude that rigorous benchmarking and assessment are central elements. The idea is not to reduce the risk of error, which would result in no discovery at all, but to minimize its costs. The "blind giant" metaphor suggests that it is always very difficult to assess the stab- ility and sustainability of a specialization at an early stage.

Landabaso (2014), in line with Foray and Goenaga (2013), finds a general reason for the difficulties that regions are currently facing when implementing RIS3 and it is related to the role of the government: the lack of an entrepreneurial public sector. He refers, among other issues,

> to government structures filled with economic development professionals working hand in hand with the private sector and other key players of the quadruple helix which pursue public goals in the form of economic transformation through innovation towards higher- value added markets and sustainable quality jobs. (Landabaso, 2014, p. 135)

This is something that, in Landabaso's opinion, is hard to find today except in a few devel- opment or innovation agencies, technology centres, technology parks and the like. He also refers to the need for a public sector that is able to take risks and experiment, one which is professional, accountable and works with the right system of incentives for the public good (delivery) and the need for public entrepreneurs that have the capacity to avoid some of the mistakes of the past in relation to old industrial policies (Ahner & Landabaso, 2011), including both the "dependency" inertia and private interests of some of the aforesaid local stakeholders.

This observation, detection and evaluation capabilities make it necessary for regional governments to work with sophisticated programmes and methodologies (Coffano & Foray, 2014; Foray, 2014). Kroll, Muller, Schnabl, and Zenker (2014) mention that some training and capacity building will inevitably be needed to enable administrations to implement the "new generation of policy measures". These authors conclude that in the case of Spain, for example, lack of competences for strategy building and the absence of multilevel governance between the central and regional governments lead to seeking assistance from consultants.

Going further into what Kroll (2015, p. 5) defines as "internal issues of regional govern- ance", the cases in this paper fit with his argument that political habits, practices and rou- tines (policy-making modes) are of substantial importance for the degree of efficacy and efficiency with which new initiatives such as RIS3 can be implemented. Foray (2014) underlines that path dependence cannot be avoided and that implementation is con- ditioned by the region's policy-making system. As McCann and Ortega-Argilés (2014, p. 24) state, the "smart specialization approach requires a serious reflection on a region's assets, capabilities and weaknesses". Camagni *et al.* (2014) support this when they note that regional innovation paths strongly depend on deep-rooted territorial elements, such as history, culture and learning processes (they refer to "the territorial approach to smart specialization").

Table 2 shows the main contributions presented on this section.

Table 2
Needs for learning and capacity building

Type of challenge	Problem for implementation/suggested solutions	Authors
Need for learning and capacity building	Lack of *public sector entrepreneurs*	Landabaso (2014)
	Lack of sophisticated programmes and methodologies and training and capacity building	Foray (2014), Coffano and Foray (2014), Kroll *et al.* (2014)
	Unsuitable political habits, practices and routines	Kroll (2015), Foray (2014)
	Rooted territorial elements (history, culture …)	Camagni *et al.* (2014)

Source: authors' elaboration.

2.3. The need for new governance modes

McCann and Ortega-Argilés (2014) refer to the "softer" institutional and governance chal- lenges associated with fostering knowledge dissemination and diffusion to enhance tech- nological adoption and adaptation. This is related to the weaknesses in the linkages within the innovation system (McCann & Ortega-Argilés, 2014) and the challenge of facilitating the design of such inter-organizational connections and coordinating efforts (Coffano & Foray, 2014).

Together with the implementation problem related to the lack of capacities of regional actors and administrators to design and run strategy processes, Kroll (2015) mentions the danger of atomization of regional governments and their incapacity to advance with intra- administrative negotiations between local ministries and agencies. The concept of govern- ance used in the later discussions in this paper refers to linkages in the regional innovation system, with specific emphasis on intergovernmental linkages between different govern- ment levels within a region.

As McCann and Ortega-Argilés (2014) note, when the concept of smart specialization emerged in the Knowledge for Growth expert group, there was no explicit regional or geo- graphical dimension of the concept. That came later and some of the implementation pro- blems that regions face now are a reflection of that gap. When including economic geography and spatial economics in the discussion of smart specialization, it became evident that translating the concept to a regional context was far more complex than the sectoral arguments imply. Smart specialization remained largely non-spatial (McCann & Ortega-Argilés, 2014). By integrating several regional and subregional gov- ernments with different scales of influence and different proximity to stakeholders in the policy-making process, the later discussion on governance aims at contributing to the territorial dimension of RIS3.

Boschma (2014) and Landabaso (2014) note it is worth mentioning that

> the smart specialization approach has much to gain by drawing lessons from the "constructed advantage" literature, which emphasizes the importance of policy encouraging crossovers between related industries that can provide complementary assets … taking region-specific intangible assets as a starting point … and promoting learning processes that are context specific. (p. 134)

The bottlenecks that prevent related industries in the regions from connecting and inter- acting, the lack of complementarities with other regions and the often too broad areas of specialization can also be considered implementation problems in smart specialization processes (Boschma, 2014; Iacobucci, 2014). As Iacobucci (2014) underlines, the emphasis on R&D and innovation will depend on the region's innovative performance and the general lack of key elements for smart specialization (connectedness, entrepreneurial spirit, industrial diversity, etc.) that can condition its capacity for implementation (Capello, 2014).

Previous implementation difficulties related to regional governance are represented in Table 3.

Table 3
Need for new governance modes

Type of problem	Problem for implementation/Suggested solutions	Authors
Need for new governance modes	Lack of complementarities with other regions; no analysis of relations between sectors, lack of "cross-fertilization" between technological domains	Coffano and Foray (2014), Boschma (2014), Iacobucci (2014)
	Weaknesses among the linkages within the innovation system – inter-organizational connections and coordinated efforts	McCann and Ortega-Argilés (2014)
	Incapacity to advance intra-administrative negotiations between local ministries and agencies (*Regional governments as atomic actors*)	Kroll (2015); Boschma (2014)
	Assumption that RIS3 is a spatial process	McCann and Ortega-Argilés (2014)
	The region's innovative performance (connectedness, entrepreneurial spirit, industrial diversity,...) can condition the implementation capacity; the case of technologically poor regions	Capello (2014), Camagni *et al.* (2014), Iacobucci (2014)

Source: authors' elaboration.

3. Methodology

The paper is based on four cases related to governance and learning for smart specializ- ation. The complexity in the choice of cases responds to the need to show different per- spectives of smart specialization on different government levels in order to later argue for multilevel governance.

The background case is the regional Basque Government's RIS3. The other three are subregional governments. One is a provincial government (that of Gipuzkoa), another is the Bilbao City Council and the fourth is a county development agency (that of Goierri) created by 18 town councils in order to operate in the county. Apart from the provincial council of Gipuzkoa, they are all directly addressed as RIS3 projects. The pro- vincial council had a more general perspective of constructing a new mode of governance, which aims to facilitate not only RIS3 projects, but also any collaborative project combining top-down with bottom-up approaches and focused on territorial development.

The three subregional cases are based on action research processes where the researchers got involved in dialogical processes with policy-makers in order to solve the policy- makers' challenges and at the same time generate relevant knowledge for an academic debate (Karlsen & Larrea, 2014a).

One of the authors was involved as an action researcher in each of the projects. In all subregional projects, we participated directly with the policy-makers in charge of the pro- cesses in continuous and sequential processes of reflection in action (while the action was taking place) and reflection on action (once the action had been developed). The research- ers' role was to develop the reflection processes, integrating concepts and frameworks and cogenerating new knowledge with the policy-makers that would feed the new cycle of action and reflection. In the case of the Basque Government, the role was more indirect, acting as a friendly outsider and helping one of the researchers who was directly involved in the process to reflect about her own role in the process and in the interaction with the actors in the Basque Government.

The choice for action research was the result of the researchers' aim at combining knowledge generation for the academic community with change in

policy-making. Green- wood and Levin (2007) consider action research more as a strategy for change than as a research method. The cases were presented previously in various academic publications (Estensoro, 2012, 2015; Estensoro & Larrea, 2012, 2015; Karlsen & Larrea, 2014a,

2014b). As a result of these processes Orkestra, Basque Institute of Competitiveness is becoming a research hub where different researchers and policy-makers are finding the opportunity to learn, not only from theory, but also from practice related to all of the pre- vious long-term projects which are still operating at the time of writing this article.

4. Case studies

4.1. The regional perspective: the Basque government

Following the proposal made by the commission to the regions, RIS3 has mainly been con- sidered an endeavour for regional governments, and these have been the main actors in their implementation. That is why it is necessary to refer to the Basque Government's RIS3 project as the main process of this type in the Basque Country. In a nutshell, the Basque Government has developed a strategy where three smart specialization thematic priorities have been established (advanced manufacturing, energy and biosciences – mainly human health) besides horizontal priorities and opportunity niches (Basque Gov- ernment, 2014a, 2014b). In order to proceed to the implementation of policies related to such vertical priorities, three task groups that integrate a variety of public and private actors are already operating at the time of writing this article. Comparative assessments of the Basque case show that the Basque Country is one of the regions in Spain with a long tradition in research, technological development and innovation policy and good starting conditions with regard to RIS3 (Kroll *et al.*, 2014; OECD, 2012). Morgan (2013a, 2013b, p. 22), who has played a critical role assessing the Basque Government on RIS3, argues that the Basque Government can legitimately claim that it has been build- ing up such a strategy for the past thirty years. Most academic contributions on the case share this perspective (Aranguren, Navarro, & Wilson, 2014; Aranguren & Wilson, 2013; Valdaliso, 2014). Valdaliso, Magro, Navarro, Aranguren, and Wilson (2014) conclude that, on the one hand, scientific and technological capabilities and institutions have been created and there has been evident policy learning for the design and implementation of such complex strategies and policies as RIS3, but there are still enormous inertia and resistances to new approaches.

The goal of this article is to expand on the potential role of subregional governments to complement the regional perspective, so we do not describe the regional government's project in detail, but mainly describe the steps taken to relate to subregional governments in the process.

After establishing the priorities, in 2013 a proposal was made by the regional govern- ment together with the researchers to set up two workshops in order to identify synergies between their strategies and those being defined by different subregional governments. Representatives from the three provincial governments, from the three City Councils of the capital cities and from Garapen (the association of county development agencies) were invited to participate. At the first workshop, the Basque Government reported on the main aspects of the regional RIS3. At the second, the regional government suggested that the others should prepare reports about the synergies between their strategy and the regional RIS3. This process was not prioritized in the following stages and, consequently, the cases presented in the following sections are not formally part of the RIS3 strategy of the Basque Country. But the attempt created a precedent to construct a multilevel govern- ance model for RIS3 in the future.

4.2. The provincial perspective: Gipuzkoa Sarean

In 2009, the Provincial Council of Gipuzkoa, one of the three provincial governments in the Basque Country, started an action research process to create connectedness between territorial actors (organizations) in four target groups: firms, organizations in the knowl- edge subsystem (universities, technology centres, etc.), political groups represented in the provincial government and the

civil society. The critical concept of the project was social capital. After diagnosing social capital, and before any intervention could be implemented, the 2011 elections led to a change of government and a different political party reached the government. Almost one year before the elections, the politician who had led the first stage of the project put its results in a nutshell when he said:

> The project has been a good experience for all participants because it has been a way to meet each other and to learn how we can work with each other, and this is a way to develop social capital too. (Meeting of the executive board of Gipuzkoa Sarean on the 30 June 2010)

The new ruling party won the elections by emphasizing a shift towards more participatory approaches to policy-making. The critical concept of the project evolved from social capital to territorial development. The government, together with the researchers, defined territory and territorial development in the context of Gipuzkoa Sarean as follows:

> Territory: the actors that live in a place, with their social, economic and political organisation, their culture and institutions as well as the physical environment they are part of. Territorial development: the process of mobilisation and participation of different actors (public and private) in which they discuss and agree on the strategies that can guide individual as well as collective behaviour. (Proposal of a New Approach to Territorial Devel- opment in Gipuzkoa, Working Paper in Gipuzkoa Sarean, June 2013)

The decision was made to develop a participatory approach to territorial development with the 11 county development agencies operating in Gipuzkoa. Counties are supra-municipal and sub-provincial territorial units. They do not have a corresponding government level, but they do have county development agencies, created by municipal governments that get together in order to gain critical mass to operate development policies.

One of the government's critical decisions in the process of developing the bottom-up approach to working with development agencies was to work with an emergent concept of strategy based on learning, negotiation and collaboration and give up the traditional plan- ning approach. They were harshly criticized in the media for not having a plan, but they decided to continue with this strategy. Action research was at the core of learning, nego- tiation and collaboration.

As a result of the process, several task forces were created among representatives of the provincial government and county development agencies. One of these was focused on the energy sector. The result of learning, negotiation and collaboration was a government decree whereby counties could apply for money to hire new staff to contact firms in energy and related sectors in the county. Each of them had to make a diagnosis of the energy value chain in the county and the gaps they had and search for opportunities for new activities. The goal was to help firms that were not in the energy value chain to enter it and those who were already in the value chain to evolve towards more value- added activities. The learning approach was critical in this group as participants from every county development agency together with the representative of the provincial government followed an action research programme in which training based on the real chal- lenges of the process was given on concepts and frameworks such as value chains, clusters and the energy sector, and they were supported with the development of their own diag- nosis, reflection and action cycles. When asked what this process offered to the provincial government, the politician most closely involved in the process answered:

> This is a decision making tool for us, which helps us to decide what to do. It also offers more capabilities to implement, as now we all [referring to participants from county development agencies] understand the policy process. (Ugaitz Iturbe, interviewed on the 22 July 2015)

4.3. The county perspective: Goierri County

Goierri is a county located in the southern part of Gipuzkoa province in the Basque Country with ap-

proximately 42,000 inhabitants and an area of 271.3 square kilometres. Economic activity in Goierri is mainly concentrated in the industrial sector, where 45% of the working population is employed. Industrial activity is predominantly shaped by subcontracting relationships between large firms located in the area and competing in the global market and their dependent suppliers, most of which are small-to-medium enterprises (SMEs).

Goieki, the local development agency, was created by the 18 municipalities in the county to support development policies in Goierri. The project used as a case study in this paper is the Industrial Forum (IF), which was created in 2012 as a "multi-actor" space where shared strategies for industrial development could be created. The IF is formed by six of the largest firms in the county together with training centres, the county technological centre and the agency, and its aim to promote sustainable industrial devel- opment for the county led to a new mode of governance.

The action research process that started in late 2012 within this forum has resulted in the definition of several priorities for industrial development. One such priority is related to the aim to facilitate the transition of the county's industrial activity towards "advanced manufacturing" or "Industry 4.0".

> This is a challenge that we all [referring to all industrial firms] share. It is not about a buzz concept, but about a need. (Manager of a firm; Meeting of the Industrial Forum on the 17 June 2015)

Researchers contributed theoretical and conceptual frameworks related to smart special- ization and specific statistical analysis that made it possible to discuss the challenges for industrial development in the county. This contribution provoked dialogue between the participants, and the priority of supporting Industry 4.0 resulted from that dialogue and reflection. The IF considered that the county's industrial development strategy should focus on horizontal priorities that the maximum amount of firms could benefit from, rather than selecting and supporting specific industrial sectors or activities.

During this process, participants in the IF studied the Basque Government's RIS3 process and observed the linkages between what they were working on and the Basque Government's process. Several attempts were made to connect the processes, but there were no established channels to make this possible.

> Although we will continue supporting our priorities, we need to know what the [Basque] government's priorities are. Synergies between the different strategies can facilitate more support for our priorities. (President of a firm; Meeting of the Industrial Forum on the 19 November 2013)

4.4. *The metropolitan perspective: Bilbao NextLab*

Bilbao City Council initiated this project in 2013 with the aim of promoting the economic transformation of the city and preparing it for present and future challenges. These chal- lenges included issues as diverse as the exhaustion of the strategy of urban development transformation, the decline of gross domestic product per capita, the ageing of the popu- lation and poor performance in innovation. This suggests that the City Council wanted to be more proactive in economic and innovation policy, an area where the local government lacks formal competences.

From September 2013 to May 2015, the Mayor's Cabinet on the Bilbao City Council, together with Bilbao Ekintza (the county development agency) and the researchers, initiated an action research process that was mainly focused on the smart specialization strategy that the City Council wanted to promote.

At the beginning of the process, the researchers delivered reports containing an analysis that helped policy-makers to find out whether the previously set priorities were correct or not.

> We need to check our work … we need the university to evaluate our policies. (Head of Mayor's Cabinet on the Bilbao City Council; Meeting on the 9 September 2013)

This role, which was initially far from being an action research process, did however develop trust

between the researchers and politicians. The dialogue touched on more deli- cate issues than the politicians had initially expected, and the role of the researchers was not to evaluate or contrast any decisions that had already been made, but to raise questions that could facilitate the development of the specialization process that the City Council was seeking. The following are literal quotes from policy-makers that help to understand how the nature of the relationship with the researchers evolved. The final sentence is related to how they interpreted the government changeover in 2015.

> It is true that we had not thought about governance until we started this process … we initiated that reflection because of you … it was a new path. (Member of the Mayor's Cabinet on the Bilbao City Council; Meeting on the 3 October 2014)

> I think that we chose the right direction: you are learning and we are too. we have done a good job … Our role when we leave [referring to the next change in government] is to be honest, handing over all of our knowledge and that which we co-generated with you, because we believe that this is the best thing for the city. (Head of Mayor's Cabinet on the Bilbao City Council; Meeting on the 3 October 2014)

A main issue in the process was coordination between Bilbao's smart specialization strat- egy and the Basque Government's RIS3. The vertical priorities defined by the Basque Gov- ernment did not include any area prioritized by Bilbao's specialization strategy. The City Council believed that the potential of urban economies had not been considered by the regional government.

> If, after all this work, the [Basque] government considers our strategy, my aim will have been achieved. My aim is to generate an "urban lobby". (Head of Mayor's Cabinet on the Bilbao City Council; Meeting on the 16 September 2014)

Although the Basque Government's aforesaid attempt at coordination was interrupted, the Bilbao City Council and the development agency remained proactive in their attempt to argue that the specialization prioritized by the city in terms of knowledge-intensive business services (KIBS) would facilitate the specialization in "advanced manufacturing" that the regional RIS3 supported.

> For us it is very important to remember that it is not all about manufacturing and products. The integration of creative industries and advanced services also means considering the urban dimension. (Director of Bilbao Ekintza; Meeting on the 16 September 2014)

5. Concluding remarks and lessons learnt

In this section, we present the main lessons learnt from the cases presented above. The discussion does not come exclusively from either theory or practice. The action research processes presented are the result of continuously testing theories in practice and reflecting on practices that can contribute to theory.

Three critical discussions have emerged that connect to the three challenges posed in the section about problems with implementation. The first is that regions are not homo- geneous units in terms of territorial scales for the implementation of RIS3, although they have often been treated in the literature as such. This lesson is directly connected to the challenge of constructing new governance modes (see Table 3). The second is that the approach to complexity in the literature is unclear and has underestimated power issues. The integration of power issues can help understand how bottom-up processes develop (Table 1). The third is that there is an implicit approach to social research on RIS3 that positions social researchers as outside observers of these processes when they can play a role as insiders in the construction of new modes of governance. We propose action research as an approach that responds to the challenges posed in terms of learning and capacity building (Table 2). The next sections focus on these discussions.

5.1. Connecting regional and subregional governments

The case studies show that regional governments might have the competences for RIS3, but often

lack the capacity to be present in many of the spaces where opportunities can be found. Policy-makers (elected politicians, civil servants and other staff in governments and their agencies) would benefit from a long-term dialogue process with representatives of firms, technology centres, universities and other organizations with potential useful knowledge for the RIS3 strategy. Our practical experience trying to operationalize the concept of entrepreneurial discovery shows that regional governments often lack people to get involved in this dialogical process. But subregional (provincial, county, local or municipal) governments that lack the competences for RIS3 do sometimes have staff with long-term trust relationships with such stakeholders. We propose that multi-scalar governance can contribute to enhance a territory's capabilities to develop RIS3 approaches as governments in different subregional scales have close relations with a multiplicity of private actors with whom regional governments cannot directly interact. This interaction is what we referred to as the senses of regional governments: people involved in the dia- logue with stakeholders can be physically close to stakeholders to share the intangible assets in the context, to listen and to see what is going on. They also talk to stakeholders and can help integrate the voice of regional governments in these processes.

There is a systematic absence of references to subregional governments in the literature on RIS3. Although many regions are designing their own policies, decentralization in this specific sphere has often stopped at the regional level. Local governments (including pro- vincial or municipal governments) have in some cases been included as one more stake- holder in local partnerships (Ortega-Argilés, 2012); still, there is no clear role defined for them or systematic academic reflection on the role of different government levels in this type of process.

Some authors recognize interdependencies between different levels of government and consequently propose multilevel governance to enhance their coordination (Charbit, 2011; Hooghe, Marks, & Schakel, 2010; Vanthillo, Vanoutrive, & Verhetsel, 2014; Vanthillo & Verhetsel, 2012). But there is little literature on how this coordination should be constructed. The cases in this paper showed attempts to construct such processes.

Of course, we claim no representativeness for the Basque case and we are aware that the multilevel governance context will be different in every region. There is a lack of cases in the literature that offer this multilevel perspective. One exception is the process in Flan- ders, where subregional discussion has been studied explicitly (Ceuninck & Reynaert, 2011), and where there has been academic production with regard to RIS3 and the role of subregional platforms as an approach to place-based development (Vanthillo *et al.*, 2014; Vanthillo & Verhetsel, 2012). They argue that these platforms have some character- istics of the place-based approach, but they also acknowledge that the goal of more integrated subregional economic development failed due to the relative powerlessness of the platforms (Cabus, 2002) and that there is no real significant decentralization in terms of competences and resources at the subregional level (Voets & De Rynck, 2006).

5.2. *Handling complexity and conflict*

When the landscape for RIS3 includes not only regional governments, but also subregional ones, the active role of different government levels and their platforms can generate a feeling of chaos. There is often a temptation to simplify and create order and structure. For instance, the Flemish government talks about downsizing the so-called crowded government house, arguing that in previous decades too many new structures and organiz- ations were established, each of them with their own competences, financial resources and decision-making procedures (Ceuninck & Reynaert, 2011). Fragmentation is also mentioned (Flemish Government, 2014). Our argument is that in order to connect to emerging processes, entrepreneurial discovery might require a certain level of complexity and that prioritizing simplicity might be a difficult challenge in these cases.

RIS3 should be able to connect the different emerging potential entrepreneurial discov- ery processes and this might be difficult to do in very simplified

structures. Making pro- cesses simpler for the regional government might hinder potential discoveries. This is a challenge for multilevel governance and an adequate definition of roles for subregional governments could allow for creative solutions without increasing inefficiency. Based on the previous cases, we argue for the development of capabilities to handle complexity as a more efficient behaviour than avoiding complexity.

Our contribution to this debate comes from discussions held with territorial actors (mainly governments) that are working to implement RIS3 processes on subregional levels. The lessons learnt led us to go beyond an interpretation of complexity that under- lines that RIS3 is a process consisting of interwoven parts that is difficult to understand. The concept that best helped in the processes was one that introduces power issues by claiming that there is a situation of territorial complexity (Karlsen, 2010; Karlsen & Larrea, 2014a) when there are autonomous but interdependent actors involved in a process that might have different interpretations of what the challenges and the answers to such challenges are, and where none of the actors are in a position to instruct the others on how to proceed. This, of course, does not mean that no actors are more powerful than others and some of them do influence the process in very clear directions. However, when there is a situation of territorial complexity, command and control pro- cesses do not work well and alternative work methods and spaces must be constructed in order to keep dialogue going.

Dialogue is at the core of development of bottom-up processes and it requires empow- erment processes not only of entrepreneurs, as is often underlined in the literature (Coffano & Foray, 2014), but also of different levels of subregional governments. Dialogue can only work if it is assumed that although there will be powerful players, governance is not a question of the other actors (either private or other government levels) adapting to the regional government's plan. It is a question of handling a situation where every actor might have the legitimacy to propose contradictory positions on smart specialization and RIS3. The case studies have revealed such complexity.

5.3. Integrating social researchers in the process

Our third argument is that social researchers can play a more varied role than we are playing in most RIS3 processes nowadays. Although social researchers are often involved in RIS3 processes, this is in the role of experts who analyse the region, diagnose the situ- ation, help reflect on governance and priorities, make recommendations on these issues and try to measure and evaluate how governments or private actors are dealing with the process. These roles are usually played out as outsiders to the process.

When working on the concepts and methods of RIS3 with governments and agencies, there is a role seldom played by social researchers, to help construct new modes of govern- ance to handle complexity. This involves complementing the mainstream methodological approaches to research on RIS3 with other research approaches that integrate the role of social researchers as facilitators of social processes. In this paper, we presented four cases inspired by action research, an approach that may help to develop this role.

There is a wide variety of approaches to action research. The approach to the case studies was systematized in Karlsen and Larrea (2014a, 2014b) and Estensoro (2015) and is based on the principles of pragmatic action research (Greenwood & Levin, 2007; Gustavsen, 1992; Johnsen, Knudsen, & Normann, 2014). We propose action research as a strategy for change that requires the researcher to play the role of facilitator of dialogue processes. Concepts and frameworks are discussed among participants in the RIS3 process in order to build trust and a shared view that can help to handle complexity. This means that the researcher's interaction with the participants in RIS3 shifts from interviews, surveys and observation to dialogue and co-generation of knowledge and from reports and papers to collective knowing (Karlsen & Larrea, 2014a) that is expressed in action.

Our final argument in this article is that action research could be an approach to learn- ing and capability development that helps construct new governance modes for bottom-up processes in RIS3.

6. Disclosure statement

No potential conflict of interest was reported by the authors.

7. References

Ahner, D., & Landabaso, M. (2011). Regional policies in times of austerity. European Review of Industrial Economics and Policy, 2. Retrieved from http://revel.unice.fr/eriep/index.html?id=3238

Aranguren, M. J., Navarro, M., & Wilson, J. R. (2014). Constructing research and innovation strat- egies for smart specialization (RIS3): Lessons from practice in three European regions. In J. M. Valdaliso, & R. J. Wilson (Eds.), Strategies for shaping territorial competitiveness (pp. 218-242). New York: Routledge.

Aranguren, M. J., & Wilson, J. R. (2013). What can experience with clusters teach us about fostering regional smart specialization? Ekonomiaz, 83(2), 126-145.

Barca, F. (2009). An agenda for a reformed cohesion policy. A place-based approach to meeting European Union challenges and expectations. Brusels: CEC.

Basque Government. (2014a). PCTI Euskadi 2020. Research & innovation smart specialisation strat- egy-RIS3. Vitoria-Gasteiz: Author.

Basque Government. (2014b). RIS3 Euskadi. Prioridades estratégicas de especialización inteligente de Euskadi. Vitoria-Gasteiz: Author.

Boschma, R. (2014). Constructing regional advantage and smart specialisation: Comparison of two European policy concepts. Scienze Regionali-Italian Journal of Regional Science, 13(1), 51-68. doi:10.3280/SCRE2014-001004

Cabus, P. (2002). Governance in Flanders' regional policy: Sub-regional platforms as development coalitions. Belgeo, 3, 277-294.

Camagni, R., Capello, R., & Lenzi, C. (2014). A territorial taxonomy of innovative regions and the European regional policy reform: Smart innovation policies. Scienze Regionali-Italian Journal of Regional Science, 13(1), 69-106. doi:10.3280/SCRE2014-001005

Capello, R. (2014). Smart specialisation strategy and the new EU cohesion policy reform: Introductory remarks. Scienze Regionali-Italian Journal of Regional Science, 13(1), 5-15. doi:10.3280/SCRE2014-001001

Ceuninck, K., & Reynaert, H. (2011). Flanders heading towards its own state reform. HKJU – Croatian and Comparative Public Administration, 11(4), 1017-1039.

Charbit, C. (2011). Governance of public policies in decentralised contexts: The multi-level approach. OECD Regional Development Working Papers.

Coffano, M., & Foray, D. (2014). The centrality of entrepreneurial discovery in building and imple- menting a smart specialisation strategy. Scienze Regionali-Italian Journal of Regional Science, 13 (1), 33-50. doi:10.3280/SCRE2014-001003

Estensoro, M. (2012). Local networks and socially innovative territories. The case of the Basque region and Goierri county. PhD thesis. Bilbao: University of the Basque Country.

Estensoro, M. (2015). How can social innovation be facilitated? Experiences from an action research process in a local network. Systemic Practice and Action Research. doi:10.1007/s11213-015-9347-2

Estensoro, M., & Larrea, M. (2012). The evolution of local development agencies from service pro- viders to facilitators in knowledge networks in the Basque Country: The role of academic exper- tise in the change process. In N. Bellini, M. Danson, & H. Halkier (Eds.), Regional development agencies: The next generation? (pp. 226-243). New York, NY: Routledge.

Estensoro, M., & Larrea, M. (2015). Vers un nou model de desenvolupament territorial: el repte de reaprendre formes de treball (Working Paper Series). Barcelona: Pacte Industrial de la Regió Metropolitana de Barcelona.

Flemish Government. (2014). The strategic policy framework for smart specialisation in Flanders. Policy Note of the Department Economy, Science and Innovation (rev. 12/2014).

Foray, D. (2013). The economic fundamentals of smart specialisation. Ekonomiaz, 2, 54-78. Foray, D. (2014). Smart specialisation, opportunities and challenges for regional innovations policy. London: Routledge.

Foray, D., David, P. A., & Hall, B. H. (2009). Smart specialisation – The concept. Knowledge Economists Policy Brief, 9. Brussels: European Commission. Retrieved from http://ec.europa. eu/invest-in-research/pdf/download_en/kfg_policy_brief_no9.pdf.

Foray, D., David, P. A., & Hall, B. H. (2011). Smart specialization. From academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation. MTEI Working Paper 2011-01.

Foray, D., Goddard, J., Goenaga Beldarrain, X., Landabaso, M., McCann, P., Morgan, K., … Ortega- Artile, R. (2012). Guide to research and innovation strategies for smart specialisation (RIS3). European Commission, Smart Specialisation Platform. Retrieved from http://ec.europa.eu/ regional_policy/sources/docgener/presenta/smart_specialisation/smart_ris3_2012.pdf.

Foray, D., & Goenaga, X. (2013). The goals of smart specialization. RIS3 Policy Brief Series, 1. Brussels: European Commission, Joint Research Centre Scientific and Policy Reports. Retrieved from http://ftp.jrc.es/EURdoc/JRC82213.pdf.

Greenwood, D. J., & Levin, M. (2007). Introduction to action research – 2nd Edition. Thousand Oaks, CA: Sage.

Gustavsen, B. (1992). Dialogue and development. Assen-Maastricht: Van Gorcum.

Hooghe, L., Marks, G., & Schakel, A. (2010). The rise of regional authority: A comparative study of 42 democracies. London: Taylor & Francis.

Iacobucci, D. (2014). Designing and implementing a smart specialisation strategy at regional level: Some open questions. Scienze Regional – Italian Journal of Regional Science, 13(1), 107-126. doi:10.3280/SCRE2014-001006

Johnsen, H. C. G., Knudsen, J. P., & Normann, R. (2014). Action research strategies at the "third place". International Journal of Action Research, 10(2), 235-256. doi:10.1688/IJAR-2014-02-Johnsen

Karlsen, J. (2010). Regional complexity and the need for engaged governance. Ekonomiaz, 74(2-10), 90-111.

Karlsen, J., & Larrea, M. (2014a). Territorial development and action research: Innovation through dialogue. Farnham: Gower.

Karlsen, J., & Larrea, M. (2014b). The contribution of action research to policy learning: The case of Gipuzkoa Sarean. International Journal of Action Research, 10(2), 129-155. doi:10.1688/IJAR-2014-02-Karlsen

Kroll, H. (2015). Efforts to implement smart specialization in practice – Leading unlike horses to the water. European Planning Studies, 23(10), 2079-2098. doi:10.1080/09654313.2014.1003036

Kroll, H., Muller, E., Schnabl, E., & Zenker, A. (2014). From smart concept to challenging practice: How European regions deal with the commission's request for novel innovation strategies. Working Papers Firms and Regions, 2. Karlsruhe: Fraunhofer ISI. Retrieved from http://www. isi.fraunhofer.de/isi-wAssets/docs/p/de/arbpap_unternehmen_region/2014/ap_r2_2014.pdf.

Landabaso, M. (2014). Time for the real economy: The need for new forms of public entrepreneur- ship. Scienze Regionali-Italian Journal of Regional Science, 13(1), 127-140. doi:10.3280/SCRE2014-001007

McCann, P., & Ortega-Argilés, R. (2011). Smart specialisation, regional growth and applications to EU Cohesion policy. Economic Geography Working Paper, University of Groningen.

McCann, P., & Ortega-Argilés, R. (2013). Modern regional innovation policy. Cambridge Journal of Regions, Economy and Society, 6(2), 187-216. doi:10.1093/cjres/rst007

McCann, P., & Ortega-Argilés, R. (2014). The role of the smart specialisation agenda in a reformed EU cohesion policy. Scienze Regionali-Italian Journal of Regional Science, 13(1), 15-32. doi:10.3280/SCRE2014-001002

Morgan, K. (2013a). Basque Country RIS3: An expert assessment on behalf of DG regional and urban policy (unpublished document).

Morgan, K. (2013b). The regional state in the era of smart specialization. Ekonomiaz, 83(2), 102-125.

OECD. (2012). Final DRAFT synthesis report on innovation-driven growth in regions: The role of sm art specialisation. Retrieved from file:///C:/Users/mestenso/Downloads/Final_Draft_Sm Sp_0ECD_EC291112.pdf.

Ortega-Argilés, R. (2012). The transatlantic productivity gap: A survey of the main causes. Journal of Economic Surveys, 26(3), 395-419. doi:10.1111/j.1467-6419.2012.00725.x

Valdaliso, J. M. (2014). The Basque Country: Past trajectory and path dependency in policy – and strategy-making. In J. M. Valdaliso, & R. J. Wilson (Eds.), Strategies for shaping territorial com- petitiveness (pp. 113-130). New York: Routledge.

Valdaliso, J. M., Magro, E., Navarro, M., Aranguren, M. J., & Wilson, R. J. (2014). Path dependence in policies supporting smart specialisation strategies. Insights from the Basque case. European Journal of Innovation Management, 17(4), 390-408. doi:10.1108/EJIM-12-2013-0136

Vanthillo, T., Vanoutrive, T., & Verhetsel, A. (2014). Towards a more "place based" industrial policy in Flanders (Belgium)? The case of the logistics sector, Draft paper for the congress Geography of Innovation, Utrecht, 2014.

Vanthillo, T., & Verhetsel, A. (2012). Paradigm change in regional policy: Towards smart special- ization? Lessons from Flanders (Belgium). Belgian Journal of Geography, 1-2, 2-16.

Voets, J., & De Rynck, F. (2006). Rescaling territorial governance: A Flemish perspective. European Planning Studies, 14(7), 905-922. doi:10.1080/09654310500496297

Gorrotoa sarean automatikoki detektatzeko teknikak ez dira oraindik perfektuak

Ziberjazarpena eta gorrotozko diskurtsoa handitzen ari dira, eta horrek jazarpenaren aurkako politikak eskatzen ditu. Hala ere, zaila da haiek detektatzea eta ikertzea, Facebook, Twitter eta gainerako sare sozialetan eta blogetan edukiak azkar ugaritzen ari direlako. Gainera, gorrotozko diskurtsoa identifikatzea konplexua izan daiteke, hiztunak kaltea eragiteko asmoa ote duen argitu behar delako. Gorrotozko diskurtsoa automatikoki atzemateko adimen artifizialeko teknikak baliatu dituzte. Ziberjazarpenaren garrantzia gorakada ikusita, baliabide gehiago behar dira detekzio-teknikak fintzeko.

Gorrotozko diskurtsoa da norbaiti buruz gaizki esaka aritzea arrazari edo generoari lotutako ezaugarriengatik. *Stormfront* foroak, gorrotozko diskurtsoaren datu-multzo berri bat argitaratu du, ikerketari laguntzeko asmoz. Eta lan horretarako adimen artifizialeko teknikak erabili dituzte, *GitHub* erreminta ezagunari esker eskuragarri daudenak.

«Gorrotozko diskurtsoan etniari eta generoari lotutako gorrotoa dira kategoriarik ohikoenak»

10.578 esaldi aztertu dituzte. Sistemak banaka sailkatzen du esaldi bakoitza: gorrotozko diskurtsoa ote den, ez den, edo berariazko harreman-kategoria bat, non gorrotozko diskurtsoa inplizitua baitago beste esaldi batzuekin konbinatzean. Sailkapen horren gidalerroak kontu handiz prestatu ziren, idazleen arteko koherentzia bermatzeko. Gero, esaldi laburregiak edo luzeegiak kendu zituzten, datu «garbiak» sortzeko.

Gorrotoa bilatzeko bidea

Datu-multzoa desorekatuta dago: gorrotorik gabeko esaldiak ugariagoak dira gorrotoa dutenak baino. Gorroto-indize bat kalkulatu zen, gorrotozko diskurtsoari lotutako hitzak identifikatzeko eta, beraz, gorrotoaren hiztegi bat osatzeko. Gorrotozko diskurtsoaren datu-basearekin gainjarrita, etnia eta generoa dira kategoriarik ohikoenak.

Artikuluak oinarrizko esperimentuak aurkezten ditu, gorrotozko testuen datu-multzo batean eginak. Datu-multzoko esaldiak etiketatuta daude —gorrotodunak edo gorrotorik ez dutenak—, esperimentuan egin ziren oharpenen baliozkotasuna frogatzeko eta etorkizuneko ikerketetarako erreferentzia ezartzeko.

Erroreak ere aztertu egin ziren. Sistemak «gorrotorik gabe» etiketaz sailkatzen zituen zenbait esaldi, lehenago eskuz «gorrotozkoa» etiketaz sailkatutakoak. Horren arrazoia izaten zen, oro har, sistemak testuingurua falta zuela. Eta kontrako akats-mota ere izaten zen; sistemak «gorrotozkoa» sailkatzen zituen zenbait esaldi, lehenago eskuz «gorrotorik gabe» etiketaz sailkatutakoak. Arrazoia izaten zen esaldiak ohiko hiztegi iraingarria erabiltzen zuela, kalterik egiteko asmorik gabe.

Esperimentuek gorroto-adierazpenak sailkatzemetodoen erronkak nabarmendu zituzten, batez ere testuingurua eta ezagutza funtsezkoak direnean emaitza zehatzak lortzeko. Gai garrantzitsua izanik, baliabide gehiago jarri beharko dira sarean gorrotoa detektatzeko teknikak hobeak izan daitezen.

InpAkta 00 (2023) 285-297
http://dx.doi.org/10.18653/v1/W18-5102

Hate speech dataset from a white supremacy forum

Ona de Gibert, Naiara Perez, Aitor García-Pablos, Montse Cuadros
HSLT Group at Vicomtech

ABSTRACT: Hate speech is commonly defined as any communication that disparages a target group of people based on some characteristic such as race, colour, ethnicity, gender, sexual orientation, nationality, religion, or other characteristic. Due to the massive rise of user-generated web content on social media, the amount of hate speech is also steadily increasing. Over the past years, interest in online hate speech detection and, particularly, the automation of this task has continuously grown, along with the societal impact of the phenomenon. This paper describes a hate speech dataset composed of thousands of sentences manually labelled as containing hate speech or not. The sentences have been extracted from Stormfront, a white supremacist forum. A custom annotation tool has been developed to carry out the manual labelling task which, among other things, allows the annotators to choose whether to read the context of a sentence before labelling it. The paper also provides a thoughtful qualitative and quantitative study of the resulting dataset and several baseline experiments with different classification models. The dataset is publicly available.

1. Introduction

The rapid growth of content in social networks such as Facebook, Twitter and blogs, makes it impossible to monitor what is being said. The increase of cyberbullying and cyberterrorism, and the use of hate on the Internet, make the identification of hate in the web an essential ingredient for anti-bullying policies of social media, as Facebook's CEO Mark Zuckerberg recently acknowledged[1]. This paper releases a new dataset of hate speech to further investigate the problem.

Although there is no universal definition for *hate speech*, the most accepted definition is provided by

[1] https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/

* **Corresponding author / Harremanetan jartzeko:** Ona de Gibert. HSLT Group at Vicomtech, Donostia/San Sebastián, Spain. e-mail: odegibert@vicomtech.org. **Contributed by** Arantza Del Pozo Echezarreta **editoreak egindako ekarpena**.
How to cite / Nola aipatu: de Gibert, Ona *et al.* (2023). «Hate Speech Dataset from a White Supremacy Forum», *Proceedings of the Second Workshop on Abusive Language Online (ALW2)*, 2018, pp. 11-20. (http://dx.doi.org/10.18653/v1/W18-5102)

Nockleby (2000): "any communication that disparages a target group of people based on some characteristic such as race, colour, ethnicity, gender, sexual orientation, nationality, religion, or other characteristic". Consider the following[2]:

(1) "God bless them all, to hell with the blacks"

This sentence clearly contains hate speech against a target group because of their skin colour. However, the identification of hate speech is often not so straightforward. Besides defining hate speech as a verbal abuse directed to a group of people because of specific characteristics, other definitions of hate speech in previous studies care to include the speaker's determination to inflect harm (Davidson *et al.*, 2017).

In all, there seems to be a pattern shared by most of the literature consulted (Nockleby, 2000; Djuric *et al.*, 2015; Gitari *et al.*, 2015; Nobata *et al.*, 2016; Silva *et al.*, 2016; Davidson *et al.*, 2017), which would define hate speech as *a)* a deliberate attack, *b)* directed towards a specific group of people, and *c)* motivated by actual or perceived aspects that form the group's identity.

This paper presents the first public dataset of hate speech annotated on Internet forum posts in English at sentence-level. The dataset is publicly available in GitHub[3]. The source forum is Stormfront[4], the largest online community of white nationalists, characterised by pseudo-rational discussions of race (Meddaugh and Kay, 2009), which include different degrees of offensiveness. Stormfront is known as the first hate website (Schafer, 2002).

The rest of the paper is structured as follows: Section 2 describes the related work and contextualises the work presented in the paper; Section 3 introduces the task of generating a manually labelled hate speech dataset; this includes the design of the annotation guidelines, the resulting criteria, the inter-annotator agreement and a quantitative description of the resulting dataset; next, Section 4 presents several baseline experiments with different classification models using the labelled data; finally, Section 5 provides a brief discussion about the difficulties and nuances of hate speech detection, and Section 6 summarises the conclusions and future work.

2. Related Work

Research on hate speech has increased in the last years. The conducted studies are diverse and work on different datasets; there is no official corpus for the task, so usually authors collect and label their own data. For this reason, there exist few publicly available resources for hate speech detection.

Hatebase[5] is the an online repository of structured, multilingual, usage-based hate speech. Its vocabulary is classified into eight categories: archaic, class, disability, ethnicity, gender, nationality, religion, and sexual orientation. Some studies make use of Hatebase to build a classifier for hate speech (Davidson *et al.*, 2017; Serra *et al.*, 2017; Nobata *et al.*, 2016). However, Saleem *et al.* (2016) prove that keyword-based approaches succeed at identifying the topic but fail to distinguish hateful sentences from clean ones, as the same vocabulary is shared by the hateful and target community, although with different intentions.

Kaggle's Toxic Comment Classification Challenge dataset[6] consists of 150k Wikipedia comments annotated for toxic behaviour. Waseem and Hovy (2016) published a collection of 16k tweets classified into racist, sexist or neither. Sharma *et al.* (2018) collected a set of 9k tweets containing harmful speech and they manually annotated them based on their degree of hateful intent. They describe three different classes of hate speech. The definition on which this paper is based overlaps mostly with their Class I, described as speech *a)* that incites violent actions, *b)* directed at a particular group, and *c)* with the intention of conveying hurting sentiments.

[2] The examples in this work may contain offensive language. They have been taken from actual web data and by no means reflect the authors' opinion.

[3] https://github.com/aitor-garcia-p/hate-speech-dataset

[4] www.stormfront.org

[5] https://www.hatebase.org/

[6] https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/data

Google and Jigsaw developed a tool called Perspective[7] that measures the "toxicity" of comments. The tool is published as an API and gives a toxicity score between 0 and 100 using a machine learning model. Such model has been trained on thousands of comments manually labelled by a team of people[8]; to our knowledge, the resulting dataset is not publicly available.

The detection of hate speech has been tackled in three main different ways. Some studies focus on subtypes of hate speech. This is the case of Warner and Hirschberg (2012), who focus on the identification of anti-Semitic posts versus any other form of hate speech. Also in this line, Kwok and Wang (2013) target anti-black hate speech. Badjatiya *et al.* (2017); Gambäck and Sikdar (2017) study the detection of racist and sexist tweets using deep learning.

Other proposals focus on the annotation of hate speech as opposed to texts containing derogatory or offensive language (Davidson *et al.*, 2017; Malmasi and Zampieri, 2017, 2018; Watanabe *et al.*, 2018). They build multi-class classifiers with the categories "hate", "offensive", and "clean".

Finally, some studies focus on the annotation of hate speech versus clean comments that do not contain hate speech (Nobata *et al.*, 2016; Burnap and Williams, 2015; Djuric *et al.*, 2015). Gitari *et al.* (2015) follow this approach but further classify the hateful comments into two categories: "weak" and "strong" hate. Del Vigna *et al.* (2017) conduct a similar study for Italian.

In all, experts conclude that annotation of hate speech is a difficult task, mainly because of the data annotation process. Waseem (2016) conducted a study on the influence of annotator knowledge of hate speech on classifiers for hate speech. Ross *et al.* (2016) also studied the reliability of hate speech annotations and acknowledge the importance of having detailed instructions for the annotation of hate speech available.

This paper aims to tackle the inherent subjectivity and difficulty of labelling hate speech by following strict guidelines. The approach presented in this paper follows (Nobata *et al.*, 2016; Burnap and Williams, 2015; Djuric *et al.*, 2015) (i.e., "hateful" versus "clean"). Furthermore, the annotation has been performed at sentence level as opposed to full-comment annotation, with the possibility to access the original complete post for each sentence. To our knowledge, this is the first work that releases a manually labelled hate speech dataset annotated at sentence level in English posts from a white supremacy forum.

3. Hate Speech Dataset

This paper presents the first dataset of textual hate speech annotated at sentence-level. Sentence-level annotation allows to work with the minimum unit containing hate speech and reduce noise introduced by other sentences that are clean.

A total number of 10,568 sentences have been extracted from Stormfront and classified as conveying hate speech or not, and into two other auxiliary classes, as per the guidelines described in Section 3.2. In addition, the following information is also given for each sentence: a post identifier and the sentence's position in the post, a user identifier, a sub-forum identifier[9]. This information makes it possible re-build the conversations these sentences belong to. Furthermore, the number of previous posts the annotator had to read before making a decision over the category of the sentence is also given.

3.1. Data extraction and processing

The content was extracted from Stormfront using web-scraping techniques and was dumped into a database arranged by sub-forums and conversation threads (Figea *et al.*, 2016). The extracted forum

[7] https://www.perspectiveapi.com

[8] https://www.nytimes.com/2017/02/23/technology/google-jigsaw-monitor-toxic-online-comments.html

[9] All the identifiers provided are fake placeholders that facilitate understanding relations between sentences, Stormfront users, etc., but do not point back to the original source.

content was published between 2002 and 2017. The process of preparing the candidate content to be annotated was the following:

1. A subset of 22 sub-forums covering diverse topics and nationalities was random-sampled to gather individual posts uniformly distributed among sub-forums and users.
2. The sampled posts were filtered using an automatic language detector[10] to discard non-English texts.
3. The resulting posts were segmented into sentences with ixa-pipes (Agerri *et al.*, 2014).
4. The sentences were grouped forming batches of 500 complete posts (~ 1,000 sentences per batch).

The manual annotation task was divided into batches to control the process. During the annotation of the first two batches, the annotation procedure and guidelines were progressively refined and adapted. In total, 10,568 sentences contained in 10 batches have been manually annotated.

A post-processing step was performed to filter excessively long or short sentences. The cleansing process removed sentences shorter than 3 words or longer than 50 words. In total, 652 sentences were discarded, which represent 6.17% of the original dataset. The resulting dataset (henceforth referred to as the "clean" dataset) is the one described in Section 3.4 and used for experimentation in Section 4. The release of the dataset will contain both the raw sentences without any post-processing and the clean version, both annotated.

3.2. *Annotation guidelines*

Schmidt and Wiegand (2017) acknowledge that the procedure for hate speech annotation is fairly vague in previous studies, which translates into low agreement scores. In this study, all the annotators together created and discussed the guidelines to ensure all participants had the same understanding of hate speech. The final guidelines consider 4 types of sentences:

3.2.1. HATE

Sentences in this category contain hate speech. Hate speech is a

a) deliberate attack
b) directed towards a specific group of people
c) motivated by aspects of the group's identity.

The three premises must be true for a sentence to be categorized as HATE. Consider the following examples:

(2) "Poor white kids being forced to treat apes and parasites as their equals."
(3) "Islam is a false religion however unlike some other false religions it is crude and appeals to crude people such as arabs."

In (2), the speaker uses "apes" and "parasites" to refer to children of dark skin and implies they are not equal to "white kids". That is, it is an attack to the group composed of children of dark skin based on an identifying characteristic, namely, their skin colour. Thus, all the premises are true and (2) is a valid example of HATE. Example (3) brands all people of Arab origin as crude. That is, it attacks the group composed of Arab people based on their origin. Thus, all the premises are true and (3) is a valid example of HATE.

3.2.2. NOHATE

This label is used to categorise sentences that do not convey hate speech per the established definition. Consider the following examples:

(4) "Where can I find NS speeches and music, also historical, in mp3 format for free download on the net."
(5) "I know of Chris Rock and subsequently have hated him for a long time."

Example (4) mentions National Socialism ("NS"), but the user is just interested in documentation about it. Therefore, the sentence itself is not an attack, i.e., premise a) is not true, despite the sound assumption that the speaker forms part of a hating community. Thus, (4) is not a valid instance of HATE. Example (5) is directed towards an individu-

[10] https://github.com/shuyo/language-detection/blob/wiki/ProjectHome.md

al; thus, premise b) is false and the sentence is not a valid example of HATE, despite the sound assumption that the attack towards the individual is based on his skin colour.

Finally, it must be emphasized that the presence of pejorative language in a sentence cannot systematically be considered sufficient evidence to confirm the existence of hate speech. The use of "fag" in the following sentence:

(6) "Two black fag's holding hands."

cannot be said to be a deliberate attack, taken without any more context, despite it likely being offensive. Therefore, it cannot be considered H AT E.

3.2.3. Relation

When (6) (repeated as (7.1)) is read in context:

(7.1) "Two black fag's holding hands."
(7.2) "That's Great!"
(7.3) "That's 2 blacks won't be having kids."

it clearly conveys hate speech. The author is celebrating that two people belonging to the black minority will not be having children, which is a deliberate attack on a group of people based on an identifying characteristic. The annotation at sentence-level fails to discern that there exists hate speech in this example. The label RELATION is for specific cases such as this, where the sentences in a post do not contain hate speech on their own, but the combination of several sentences does. Consider another example:

(8.1) "Probably the most disgusting thing I've seen in the last year."
(8.2) "She looks like she has some African blood in her, or maybe it's just the makeup."
(8.3) "This is just so wrong."

Each sentence in isolation does not convey hate speech: in (8.1) and (8.3), a negative attitude is perceived, but it is unknown whether it is targeted towards a group of people; in (8.2), there is no hint of an attack, not even of a negative attitude. However, the three sentences together suggest that having "African blood" makes a situation (whatever "this" refers to) disgusting, which constitutes hate speech according to the definition proposed.

The label RELATION is given separately to all the sentences that need each other to be understood as hate speech. That is, consecutive sentences with this label convey hate speech but depend on each other to be correctly interpreted.

3.2.4. Skip

Sentences that are not written in English or that do not contain information as to be classified into HATE or NOHATE are given this label.

(9) "Myndighetene vurderer n om de skal f permanent oppholdstillatelse."
(10) "YouTube - Broadcast Yourself."

Example (9) is in Norwegian and (10) is irrelevant both for HATE and NOHATE.

3.3. Annotation procedure

In order to develop the annotation guidelines, a draft was first written based on previous similar work. Three of the authors annotated a 1,144-sentence batch of the dataset following the draft, containing only the categories HATE, NOHATE and SKIP. Then, they discussed the annotations and modified the draft accordingly, which resulted in the guidelines presented in the previous section, including the RELATION category. Finally, a different batch of 1,018 sentences was annotated by the same three authors adhering to the new guidelines in order to calculate the inter-annotator agreement.

Table 1 shows the agreements obtained in terms of the average percent agreement (*avg %*), average Cohen's kappa coefficient (Cohen, 1960) (*avg k*), and Fleiss' kappa coefficient (Fleiss, 1971) (*fleiss*). The number of annotated sentences (# sent) and the number of categories to label (# cat) are also given for each batch. The results are in line with similar works (Nobata *et al.*, 2016; Warner and Hirschberg, 2012).

Table 1
Inter-annotator agreements on batches 1 and 2

	# sent	# cat	*avg %*	*avg k*	*fleiss*
1	1,144	3	91.03	0.614	0.607
2	1,018	4	90.97	0.627	0.632

All the annotation work was carried out using a web-based tool developed by the authors for this purpose. The tool displays all the sentences belonging to the same post at the same time, giving the annotator a better understanding of the post's author's intention. If the complete post is deemed insufficient by the annotator to categorize a sentence, the tool can show previous posts to which the problematic post is answering, on demand, up to the first post in the thread and its title. This consumption of context is registered automatically by the tool for further treatment of the collected data.

As stated by other studies, context appears to be of great importance when annotating hate speech (Watanabe *et al.*, 2018). Schmidt and Wiegand (2017) acknowledge that whether a message contains hate speech or not can depend solely on the context, and thus encourage the inclusion of extra-linguistic features for annotation of hate speech. Moreover, Sharma *et al.* (2018) claim that context is essential to understand the speaker's intention.

3.4. Dataset statistics

This section provides a quantitative description and statistical analysis of the clean dataset published. Table 2 shows the distribution of the sentences over categories. The dataset is unbalanced as there exist many more sentences not conveying hate speech than "hateful" ones.

Table 3 refers to the subset of sentences that have required reading additional context (i.e. previous comments to the one being annotated) to make an informed decision by the human annotators. The category HATE is the one that requires more context, usually due to the use of slang unknown to the annotator or because the annotator needed to find out the actual target of an offensive mention.

Table 2
Distribution of sentences over categories in the clean dataset

Assigned label	# sent	%
HATE	1,119	11.29
NOHATE	8,537	86.09
RELATION	168	1.69
SKIP	92	0.93
total	9,916	100.00

Table 3
Percentage of sentences for which the human annotators asked for additional context

	Context used	No context used
HATE	22.70	77.30
NOHATE	8.00	92.00

The remaining of the section focuses only on the subset of the dataset composed of the categories HATE and NOHATE, which are the core of this work. Table 4 shows the size of said subset, along with the average sentence length for each class, their word counts and their vocabulary sizes.

Table 4
Size of the categories HATE and NOHATE in the clean dataset

	Hate	noHate
sentences	1,119	8,537
sentence length	20.39 ±9.46	15.15 ±9.16
word count	24,867	144,353
vocabulary	4,148	13,154

Regarding the distribution of sentences over Stormfront accounts, the dataset is balanced as

there is no account that contributes notably more than any other: the average percentage of sentences is of 0.50 ± 0.42 per account, the total amount of accounts in the dataset being 2,723. The sub-forums that contain more HATE belong to the category of news, discussion of views, politics, philosophy, as well as to specific countries (i.e., Ireland, Britain, and Canada). In contrast, the sub-forums that contain more NOHATE sentences are about education and homeschooling, gatherings, and youth issues.

In order to obtain a more qualitative insight of the dataset, a HATE score (H S) has been calculated based on the Pointwise Mutual Information (PMI) value for each word towards the categories HATE and NOHATE. PMI allows calculating the correlation of each word with respect to each category. The difference of the PMI value of a word *w* and the category H AT E and the PMI of the same word *w* and the category NOHATE results in the HATE score of *w*, as shown in Formula 1.

$$\text{H S}(w) = PMI\,(w, \text{H AT E}) - PMI\,(w, \text{N O HAT E}) \qquad (1)$$

Intuitively, this score is a simple way of capturing whether the presence of a word in a HATE context occurs significantly more often than in a NOHATE context. Table 5 shows the 15 most and least hateful words: the more positive a HATE score, the more hateful a word, and vice versa.

The results show that the most hateful words are derogatory or refer to targeted groups of hate speech. On the other hand, the least hateful words are neutral in this regard and belong to the semantic fields of Internet, or temporal expressions, among others. This shows that the vocabulary is discernible by category, which in turn suggests that the annotation and guidelines are sound.

Performing the same calculation with bi-grams yields expressions such as "gene pool", "race traitor", and "white guilt" for the most hateful category, which appear to be concepts related to race issues. The less hateful terms are expressions such as "white power", "white nationalism" and "pro white", which clearly state the right-wing extremist politics of the forum users.

Table 5
Most (positive HS) and least (negative HS) hateful words

	H S		H S
ape	6.81	pm	−3.38
ape	6.81	pm	−3.38
scum	6.25	group	−3.34
savages	5.73	week	−3.13
filthy	5.73	idea	−2.70
mud	5.31	thread	−2.68
homosexuals	5.31	german	−2.67
filth	5.19	videos	−2.67
apes	5.05	night	−2.63
beasts	5.05	happy	−2.63
homosexual	5.05	join	−2.63
threat	5.05	pictures	−2.60
monkey	5.05	eyes	−2.54
libtard	5.05	french	−2.52
coon	5.05	information	−2.44
niglet	4.73	band	−2.44

Finally, the dataset has been contrasted against the English vocabulary in Hatebase. 9.28% of HATE vocabulary overlaps with Hatebase, a higher percentage than for NOHATE vocabulary, of which 6.57% of the words can be found in Hatebase. In Table 6, the distribution of HATE vocabulary is shown over Hatebase's 8 categories. Although some percentages are not high, all 8 categories are present in the corpus. Most of the HATE words from the dataset belong to ethnicity, followed by gender. This is in agreement with Silva *et al.* (2016), who conducted a study to analyse the targets of hate in social networks and showed that hate based on race was the most common.

Table 6
Distribution of HATE vocabulary over Hate-base categories

category	%	examples
archaic	2.46	div, wigger
ethnicity	41.63	coon, paki
nationality	7.03	guinea, leprechaun
religion	1.34	holohoax, prod
gender	36.05	bird, dyke
sexual orientation	2.34	fag, queer
disability	2.01	mongol, retarded
social class	7.14	slag, trash
total	100.00	

4. Experiments

In order to further inspect the resulting dataset whether the two annotated classes are separable based solely on the text of the labelled instances) a set of baseline experiments have been conducted. These experiments do not exploit any external resource such as lexicons, heuristics or rules. The experiments just use the provided dataset and well-known approaches from the literature to provide a baseline for further research and improvement in the future.

4.1. Experimental setting

The experiments are based on a balanced subset of labelled sentences. All the sentences labelled as HATE have been collected, and an equivalent number of NOHATE sentences have been randomly sampled, summing up 2k labelled sentences. From this amount, the 80% has been used for training and the remaining 20% for testing.

The evaluated algorithms are the following:

- Support Vector Machines (SVM) (Hearst *et al.*, 1998) over Bag-of-Words vectors. Word-count-based vectors have been computed and fed into a Python Scikit-learn LinearSVM[11] classifier to separate HATE and NOHATE instances.
- Convolutional Neural Networks (CNN), as described in (Kim, 2014). The implementation is a simplified version using a single input channel of randomly initialized word embeddings[12].
- Recurrent Neural Networks with Long Short-term Memories (LSTM) (Hochreiter and Schmidhuber, 1997). A LSTM layer of size 128 over word embeddings of size 300.

All the hyperparameters are left to the usual values reported in the literature (Greff *et al.*, 2017). No hyperparameter tuning has been performed. A more comprehensive experimentation and research has been left for future work.

4.2. Results

The baseline experiments include a majority class baseline showing the balance between the two classes in the test set. The results are given in terms of accuracy for HATE and NOHATE individually, and the overall accuracy, calculated according to the equations 2, 3 and 4, where *TP* are the true positives and *FP* are the false positives.

$$Acc_{\text{HATE}} = \frac{TP_{\text{HATE}}}{TP_{\text{HATE}} + FP_{\text{HATE}}} \quad (2)$$

$$Acc_{\text{NOHATE}} = \frac{TP_{\text{NOHATE}}}{TP_{\text{NOHATE}} + FP_{\text{NOHATE}}} \quad (3)$$

$$Acc_{\text{ALL}} = \frac{TP_{\text{ALL}}}{TP_{\text{ALL}} + FP_{\text{ALL}}} \quad (4)$$

We show the accuracy for the both complementary classes instead of the precision-recall of a single class to highlight the performance of the classifiers for the both classes individually. Table 7 shows the results of using only sentences that did *not* require additional context to be labelled, while Table 8 shows the results of including those sentences that required additional context. Not surprisingly, the results are lower when including sentences that required additional context. If a human annotator required additional information to make

[11] http://scikit-learn.org/stable/modules/svm.html

[12] https://github.com/dennybritz/cnn-text-classification-tf

a decision, it is to expect that an automatic classifier would not have enough information or would have a harder time making a correct prediction. The results also show that NOHATE sentences are more accurately classified than HATE sentences. Overall, the LSTM-based classifier obtains better results, but even the simple SVM using bag-of-words vectors is capable of discriminating the classes reasonably well.

Table 7
Results excluding sentences that required additional context for manual annotation

	Acc_{HATE}	Acc_{NOHATE}	Acc_{ALL}
Majority	n/a	n/a	0.50
SVM	0.72	0.76	0.74
CNN	0.54	0.86	0.70
LSTM	0.76	0.80	0.78

Table 8
Results including sentences that required additional context for manual annotation

	Acc_{HATE}	Acc_{NOHATE}	Acc_{ALL}
Majority	n/a	n/a	0.50
SVM	0.69	0.73	0.71
CNN	0.55	0.79	0.66
LSTM	0.71	0.75	0.73

4.3. Error Analysis

In order to get a deeper understanding of the performance of the classifiers trained, a manual inspection has been performed on a set of erroneously classified sentences. Two main types of errors have been identified:

Type I errors Sentences manually annotated as HATE but classified as NOHATE by the system, usually due to a lack of context or to a lack of the necessary world knowledge to understand the meaning of the sentence:

(11) "Indeed, now they just need to feed themselves, educate themselves, police themselves ad nauseum…"
(12) "If you search around you can probably find 'hoax of the 20th century' for free on the net."

In (11), it is not clear without additional context who "themselves" are. It actually refers to people of African origin. In its original context, the author was implying that they are not able to feed, police nor educate themselves. This would make the sentence an example of hate speech, but it could also be a harmless comment given the appropriate context. In (12), the lack of world knowledge about what the Holocaust is, or what naming it "hoax" implies —i.e., denying its existence—, would make it difficult to understand the sentence as an act of hate speech.

Type II errors Sentences manually labelled as NOHATE and automatically classified as HATE, usually due to the use of common offensive vocabulary with non-hateful intent:

(13) "I dont like reporting people but the last thing I will do is tolerate some stupid pig who claims Hungarians are Tartars."
(14) "More black-on-white crime: YouTube - Black Students Attack White Man For Eating Dinner With Black"

In (13), the user accuses and insults a particular individual. Example (14) is providing information on a reported crime. Although vocabulary of targeted groups is used in both cases (i.e., "Hungarians", "Tartars", "black"), the sentences do not contain HATE.

5. Discussion

There are several aspects of the introduced dataset, and hate speech annotation in general, that deserve a special remark and discussion.

First, the source of the content used to obtain the resulting dataset is on its own a source of offensive language. Being Stormfront a white supremacists' forum, almost every single comment contains

some sort of intrinsic racism and other hints of hate. However, not every expression that contains a racist cue can be directly taken as hate speech. This is a truly subjective debate related to topics such as free speech, tolerance and civics. That is one of the main reasons why this paper carefully describes the annotation criteria for what here counts as hate speech and what not. In any case, despite the efforts to make the annotation guidelines as clear, rational and comprehensive as possible, the annotation process has been admittedly demanding and far from straightforward.

In fact, the annotation guidelines were crafted in several steps, first paying attention to what the literature points about hate speech annotation. After a first round of manual labelling, inconsistencies among the human annotators were discussed and the guidelines and examples were adapted. From those debates we extract some conclusions and pose several open questions. The first annotation criteria (hate speech being a *deliberate attack*) still lacks robustness and a proper definition, becoming ambiguous and subject to different interpretations. A more precise definition of what an *attack* is and what it is not would be necessary: Can an objective fact that however undermines the honour of a group of people be considered an attack? Is the mere use of certain vocabulary (e.g. "nigger") automatically considered an attack? With regard to the second annotation criteria (hate speech being *directed towards a specific group of people*), it was controversial among the human annotators as well. Sentences were found that attacked individuals and mentioned the targets' skin colour or religion, political trends, and so on. Some annotators interpreted these as indirect attacks towards the collectivity of people that share the mentioned characteristics.

Another relevant point is the fact that the annotation granularity is sentence level. Most, if not all, of the existing datasets label full comments. A comment might be part of a more elaborated discourse, and not every part may convey hate. It is arguable whether a comment containing a single hate-sentence can be considered "hateful" or not. The dataset released provides the full set of sentences per comment with their annotations, so each can decide how to work with it.

In addition, and related to the last point, one of the labels included for the manual labelling is RELATION. This label is meant to be used when two or more sentences need each other to be understood as hate speech, usually because one is a premise and the following is the (hateful) conclusion. This label has been seldom used.

Finally, a very important issue to consider is the need of additional context to label a sentence (i.e., the rest of the conversation or the title of the forum-thread). It can happen to human annotators and, of course, to automatic classifiers, as confirmed in the error analysis (Section 4.3). Studying context dependency to perform the labelling, it has been observed that annotators learn to distinguish hate speech more easily over time, requiring less and less context to make the annotations (see Figure 1).

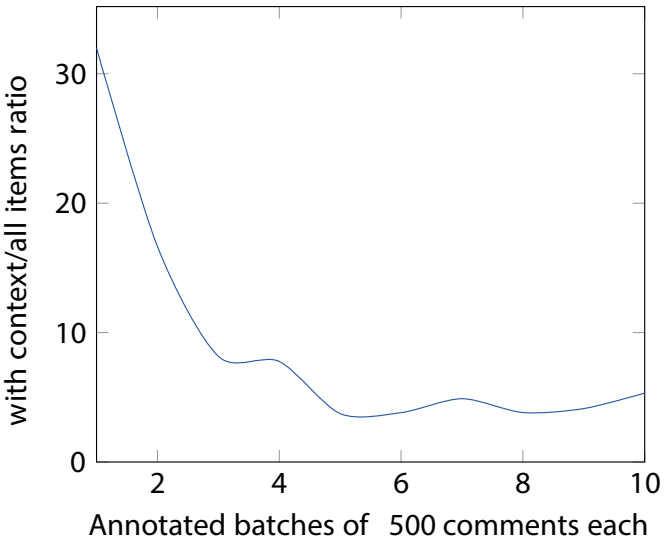


Figure 1. Percentage of comments per batch that required additional context to be manually labelled. The amount of context needed by a human annotator decreases over time

6. Conclusions and Future Work

This paper describes a manually labelled hate speech dataset obtained from Stormfront, a white supremacist online forum.

The resulting dataset consists of ~10k sentences labelled as conveying hate speech or not. Since the definition of hate speech has many subtleties, this work includes a detailed explanation of the manual annotation criteria and guidelines. Furthermore, several aspects of the resulting dataset have been studied, such as the necessity of additional context by the annotators to make a decision, or the distribution of the vocabulary used in the examples labelled as hate speech. In addition, several baseline experiments have been conducted using automatic classifiers, with a focus on examples that are difficult for automatic classifiers, such as those that required additional context or world knowledge. The resulting dataset is publicly available.

This dataset provides a good starting point for discussion and further research. As future work, it would be interesting to study how to include world knowledge and/or the context of an online conversation (i.e. previous and following messages, forum thread title, and so on) in order to obtain more robust hate speech automatic classifiers. Future studies could also explore how sentences labelled as RELATION affect classification, as this sentences have not been included in the experiments presented. In addition, more studies should be performed to characterize the content of the dataset in depth, regarding timelines, user behaviour and hate speech targets, for instance. Finally, since the proportion of HATE/NO-HATE examples tends to be unbalanced, a more sophisticated manually labelling system with active learning paradigms would greatly benefit future labelling efforts.

7. Acknowledgements

This work has been supported by the European Commission under the project ASGARD (700381, H2020-FCT-2015). We thank the Hate-base team, in particular Hatebase developer Timothy Quinn, for providing Hatebase's English vocabulary dump to conduct this study. Finally, we would like to thank the reviewers of the paper for their thorough work and valuable suggestions.

8. References

R. Agerri, J. Bermudez, and G. Rigau. 2014. IXA pipeline: Efficient and Ready to Use Multilingual NLP tools. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 3823-3828.

P. Badjatiya, S. Gupta, M. Gupta, and V. Varma. 2017. Deep Learning for Hate Speech Detection in Tweets. In *Proceedings of the 26th International Conference on World Wide Web (WWW 2017)*, pages 759-760.

P. Burnap and M. L. Williams. 2015. Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, 7(2):223-242.

J. Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37-46.

T. Davidson, D. Warmsley, M. Macy, and I. Weber. 2017. Automated Hate Speech Detection and the Problem of Offensive Language. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, pages 512-515.

F. Del Vigna, A. Cimino, F. Dell'Orletta, M. Petrocchi, and M. Tesconi. 2017. Hate me, hate me not: Hate speech detection on Facebook. In *Proceedings of the First Italian Conference on Cybersecurity (ITASEC17)*, pages 86-95.

N. Djuric, J. Zhou, R. Morris, M. Grbovic, V. Radosavljevic, and N. Bhamidipati. 2015. Hate Speech Detection with Comment Embeddings, pages 29-30.

L. Figea, L. Kaati, and R. Scrivens. 2016. Measuring online affects in a white supremacy forum. *IEEE International Conference on Intelligence and Security Informatics: Cybersecurity and Big Data, ISI 2016*, pages 85-90.

J. L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378.

B. Gamba¨ck and U. K. Sikdar. 2017. Using convolutional neural networks to classify hate-speech. In *Proceedings of the First Workshop on Abusive Language Online (ALW1)*, pages 85-90.

N. D. Gitari, Z. Zuping, H. Damien, and J. Long. 2015. A Lexicon-based Approach for Hate Speech Detection. *International Journal of Multimedia and Ubiquitous Engineering*, 10(4):215-230.

K. Greff, R. K Srivastava, J. Koutn´ık, B. R. Steunebrink, and J. Schmidhuber. 2017. LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222-2232.

M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf. 1998. Support vector machines. *IEEE Intelligent Systems and their Applications*, 13(4):18-28.

S. Hochreiter and J. Schmidhuber. 1997. Long Short-term Memory. *Neural Computation*, 9(8):1735-1780.

Y. Kim. 2014. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, pages 1746-1751.

I. Kwok and Y. Wang. 2013. Locate the Hate: Detecting Tweets against Blacks. *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, pages 1621-1622.

S. Malmasi and M. Zampieri. 2017. Detecting Hate Speech in Social Media. In *Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP 2017)*, pages 467-472.

S. Malmasi and M. Zampieri. 2018. Challenges in discriminating profanity from hate speech. *Journal of Experimental & Theoretical Artificial Intelligence*, 30(2):187-202.

P. M. Meddaugh and J. Kay. 2009. Hate Speech or "Reasonable Racism?" The Other in Stormfront. *Journal of Mass Media Ethics*, 24(4):251-268.

C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad, and Y. Chang. 2016. Abusive Language Detection in Online User Content. In *Proceedings of the 25th International Conference on World Wide Web (WWW 2016)*, pages 145-153.

J. T. Nockleby. 2000. Hate speech. *Encyclopedia of the American Constitution*, 3:1277-79.

B. Ross, M. Rist, G. Carbonell, B. Cabrera, N. Kurowsky, and M. Wojatzki. 2016. Measuring the reliability of hate speech annotations: The case of the european refugee crisis. In *Proceedings of NLP4CMC III: 3rd Workshop on Natural Language Processing for Computer-Mediated Communication*, volume 17, pages 6-9.

H. M. Saleem, K. P. Dillon, S. Benesch, and D. Ruths. 2016. A Web of Hate: Tackling Hateful Speech in Online Social Spaces. In *Proceedings of the First Workshop on Text Analytics for Cybersecurity and Online Safety (TA-COS 2016)*.

Joseph A Schafer. 2002. Spinning the web of hate: Web-based hate propagation by extremist organizations. *Journal of Criminal Justice and Popular Culture*, 9(2):69-88.

A. Schmidt and M. Wiegand. 2017. A Survey on Hate Speech Detection using Natural Language Processing. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media (SocialNLP 2017)*, pages 1-10.

J. Serra, I. Leontiadis, D. Spathis, J. Blackburn, G. Stringhini, and A. Vakali. 2017. ¡Class-based prediction errors to detect hate speech with out-of- vocabulary words. In *Abusive Language Workshop*, volume 1. Abusive Language Workshop.

S. Sharma, S. Agrawal, and M. Shrivastava. 2018. Degree based Classification of Harmful Speech using Twitter Data. *arXiv:1806.04197*.

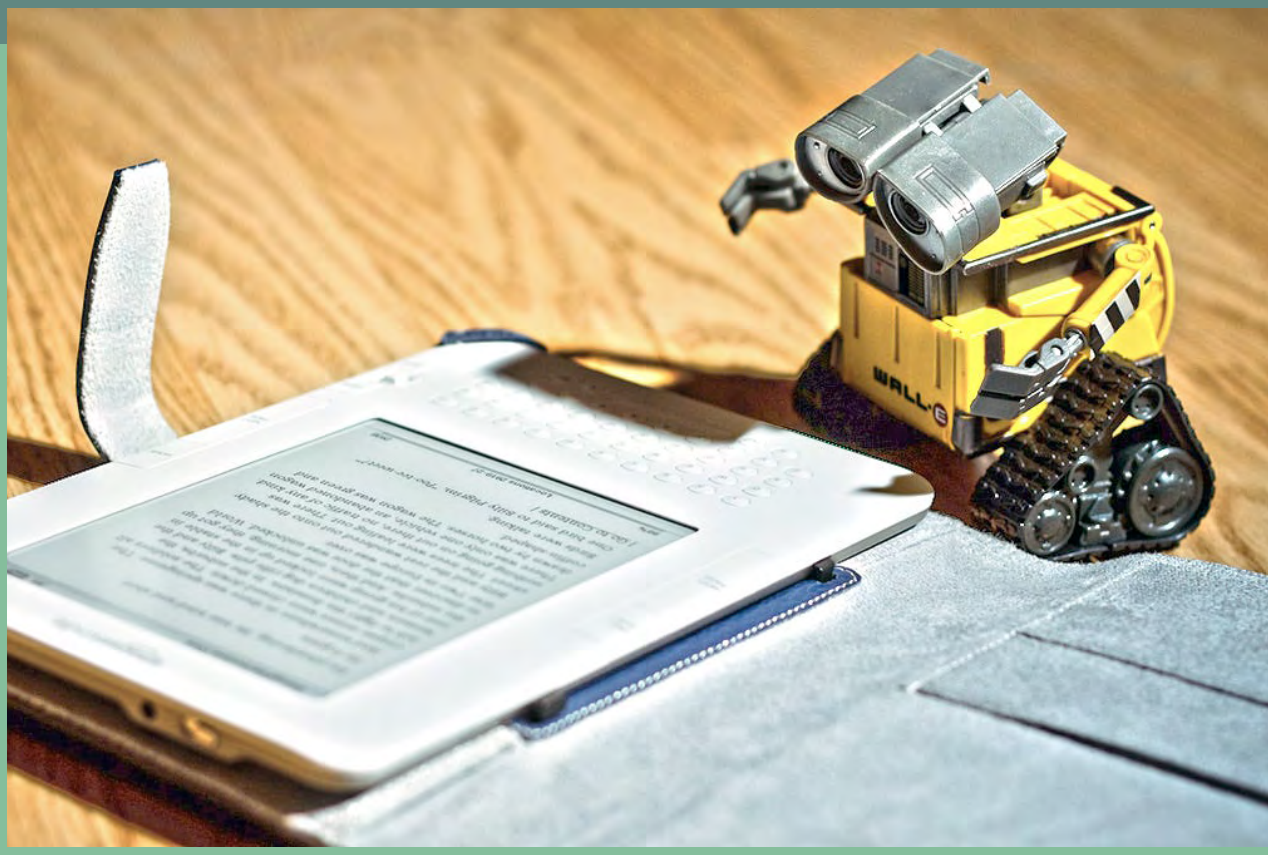
L. Silva, M. Mondal, D. Correa, F. Benevenuto, and I. Weber. 2016. Analyzing the Targets of Hate in Online Social Media. In *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*, pages 687-690.

W. Warner and J. Hirschberg. 2012. Detecting Hate Speech on the World Wide Web. In *Proceedings of the Second Workshop on Language in Social Media (LSM 2012)*, pages 19-26.

Z. Waseem. 2016. Are You a Racist or Am I Seeing Things? Annotator Influence on Hate Speech Detection on Twitter. pages 138-142.

Z. Waseem and D. Hovy. 2016. Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. In *Proceedings of the NAACL Student Research Workshop (NAACL SRW 2018)*, pages 88-93.

H. Watanabe, M. Bouazizi, and T. Ohtsuki. 2018. Hate Speech on Twitter: A Pragmatic Approach to Collect Hateful and Offensive Expressions and Perform Hate Speech Detection. *IEEE Access*, 6:13825-13835.



Robot «reading»: CC 4.0 image from futurity.org

Corpus bat txikia denean, kalitate altukoa ez izateak ez du asko eragiten

Lengoaia naturala ulertzen duten adimen artifizialeko sistemak corpusetatik abiatuta trebatzen dira; alegia, testu-multzoetatik abiatuta. Oro har, corpus horiek zenbat eta handiagoak eta kalitate altuagokoak izan, orduan eta hobea izango da trebatutako sistema. Hizkuntza txiki askotan, ordea, corpusak ez dira handiak izaten, eta batzuetan ez dira kalitate oso altukoak, gainera. Kasu horietarako, zenbat eragiten du kalitateak? Ikerketa honek erantzun dio galdera horri, euskararen kasutik abiatuta.

Itzulpen automatikoak egiteko, edo lengoaia naturala ulertu behar duten beste aplikazio batzuetarako, oso garrantzitsua da sistemak zerekin trebatzen diren. Ikertzaileek ikusi dute ingelesezkoak ez diren datu-multzoetarako iturri primarioak, hala nola *CommonCrawl*, kalitate eztabaidagarrikoak izaten direla batzuetan. Webetatik automatikoki eskuratzen dituzten testuen bildumak dira. Bilduma horien kalitateak errendimenduan eragin dezakeen galera hobeto ulertzeko, azterlanak puntu fokal gisa hartu zuen euskarazko irudikapenaren ikaskuntza.

Horretarako, *CommonCrawlen* automatikoki iragazitako datuak erabili ordez, ikuspegi pertsonalizatua probatzea erabaki zuten ikertzaileek. Eskuz identifikatu eta erauzi zituzten kalitate handiko edukiagatik ezagunak diren zenbait webgune. Datuen multzoak, *EusCrawl* izenekoak, 12,5 milioi dokumentu ditu; 33 webgunetakoak dira, denak *Creative Commons* lizentziadunak. Corpus berri hori ezagunak diren beste hizkuntza anitzeko datu-multzoen antzekoa da tamainan. Hala ere, *EusCrawl* askoz kalitate handiagokotzat jo zuten euskal hiztunek: dokumentuen % 66 kalitate onekotzat jo ziren; beste corpusen dokumentuen kasuan, berriz, % 33k baino gutxiagok jaso zituzten kalifikazio horiek.

Kalitatea ez da muga bat

Bitxia bada ere, itxurazko kalitate-desberdintasunarekin ere, lengoaia naturalaren ulermenean egindako lana ia berdina izan zen, entrenamendu aurreko zein datu-multzo erabili zen kontuan hartu gabe. Horrek iradokitzen duenez, euskara bezalako baliabide gutxiko hizkuntzak direnean, datuen kalitatea ez litzateke izango muga nagusia. Aldiz, erabakigarriagoak izan daitezke lanean beste faktore batzuk, hala nola datuen bolumena eta haien domeinu-estaldura.

> ***«Onuragarriagoa izan liteke testu-multzo zabalagoak eta askotarikoagoak biltzea, kalitate hobea dutenak bilatu ordez»***

Azken finean, baliabide gutxiko hizkuntzetan, *CommonCrawletik* eratorritako datu-multzoekin kalitate-arazo nabarmenak egon arren, litekeena da arazo horiek eragin handirik ez izatea lengoaia naturala ulertzeko zereginetan. Etorkizunerako, badirudi onuragarriagoa izango dela datu-multzo zabalagoak eta askotarikoagoak biltzea, ahalegina dauden datuen kalitatea hobetzera bideratu ordez. Gainera, hizkuntzen arteko transferentzia-metodoen potentziala nabarmentzen du ikerketak, datu eleaniztunak eraginkortasunez aprobetxatzeko. Hala ere, kontuan hartu behar da ohar horiek euskararako atazetan oinarritzen direla, eta litekeena dela emaitzak desberdinak izatea beste hizkuntza batzuetan edo bestelako zereginetan.

Ikerketak erreferentzia-puntu berri bat ezarri du datu publikoetatik abiatuta, eta, kalitate handiko *EusCrawl* corpusaren bidez, ikertzaileek ikerketa irekiago eta errepikagarriago bat sustatu nahi dute euskararen esparruan.

InpAkta 00 (2023) 298-307

http://dx.doi.org/10.18653/v1/2022.emnlp-main.499

Does corpus quality really matter for low-resource languages?

Mikel Artetxe[1*], Itziar Aldabe[2], Rodrigo Agerri[2], Olatz Perez-de-Viñaspre[2], Aitor Soroa[2]

[1]Meta AI; [2]HiTZ Center, University of the Basque Country (UPV/EHU)

ABSTRACT: The vast majority of non-English corpora are derived from automatically filtered versions of CommonCrawl. While prior work has identified major issues on the quality of these datasets (Kreutzer *et al.*, 2021), it is not clear how this impacts downstream performance. Taking representation learning in Basque as a case study, we explore tailored crawling —manually identifying and scraping websites with high-quality content— as an alternative to filtering CommonCrawl. Our new corpus, called EusCrawl, is similar in size to the Basque portion of popular multilingual corpora like CC100 and mC4, yet it has a much higher quality according to native annotators. For instance, 66% of documents are rated as high-quality for EusCrawl, in contrast with < 33% for both mC4 and CC100. Nevertheless, we obtain similar results on down-stream NLU tasks regardless of the corpus used for pre-training. Our work suggests that NLU performance in low-resource languages is not primarily constrained by the quality of the data, and other factors like corpus size and domain coverage can play a more important role.

1. Introduction

Large-scale pre-training has resulted in a paradigm shift in NLP (Bommasani *et al.*, 2021). While recent progress has been primarily driven by scaling up on model size and compute, both data quantity and quality have been shown to play a critical role (Kaplan *et al.*, 2020; Rae *et al.*, 2022). Nevertheless, existing efforts on data curation have primarily focused on English, and recent work on multilingual pre-training has relied on automatically filtered versions of CommonCrawl. For instance, XLM-R was trained on CC100 (Conneau *et al.*, 2020), mT5 was trained on mC4 (Xue *et al.*, 2021), and XGLM was trained on CC100-XL (Lin *et al.*, 2021), which were all obtained by running language identification on several CommonCrawl snapshots and filtering through language-agnostic approaches. Unfortunately, Kreutzer *et al.* (2021) identified major issues

* **Corresponding author / Harremanetan jartzeko:** Mikel Artetxe. Meta AI. e-mail: artetxe@meta.com. **Contributed by** Eneko Agirre **editoreak egindako ekarpena**.

How to cite / Nola aipatu: Artetxe, Mikel *et al.* «Does Corpus Quality Really Matter for Low-Resource Languages?», *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 2022, 7383-7390 (http://dx.doi.org/10.18653/v1/2022.emnlp-main.499)

on the quality of such multilingual datasets, ranging from language identification errors to boilerplate and non-linguistic content. However, the practical impact of these issues has not been studied, and it is unclear the extent to which higher-quality data could lead to better performance in low-resource languages.

Table 1
Basque corpora used in our experiments

	Size	Tokens	Docs	Source
mC4 (Xue *et al.*, 2021)	4,387 MiB	1,004M	30,098k	Filtered CommonCrawl
CC100 (Conneau *et al.*, 2020)	2,027 MiB	416M	16,761k	Filtered CommonCrawl
Wikipedia	313 MiB	66M	2,685k	Wikipedia dump
EusCrawl (ours)	2,149 MiB	423M	12,528k	Tailored crawling (see Table 2)

We report uncompressed text size, number of SentencePiece tokens (using a 50K vocabulary learned in each corpus), and number of documents.

Table 2
Data sources used to build EusCrawl

	Size	Tokens	Docs	License	Domain
Tokikom†	784 MiB	153M	4,961k	CC-BY-SA	Local media
Berria	525 MiB	101M	2,193k	CC-BY-SA	National newspaper
Hitza‡	418 MiB	80M	2,257k	CC-BY-NC-ND	Regional newspapers
Wikipedia	313 MiB	68M	2,685k	CC-BY-SA	Encyclopedia
Argia	101 MiB	20M	370k	CC-BY-SA	News magazine
Bilbo Hiria irratia	7 MiB	1M	54k	CC-BY-NC-SA	Radio station
Sarean	2 MiB	0.3M	8k	CC-BY-SA	Technology blog

† Tokikom is a network of local media; we include Aiaraldea, Aikor, Anboto, Tolosaldeko Ataria, Aiurri, Erran, Euskalerria Irratia, Goiena, Guaixe, Hiruka, Karkara, Maxixatzen, Plaentxia, Alea, Noaua, Txintxarri, Uztarria, Amezti, Zarauzko Hitza, Kronika and Geuria. ‡ Hitza is a family of regional newspapers; we include Bidasoko Hitza, Busturialdeko Hitza, Goierriko Hitza, Irutxuloko Hitza, Lea-Artibai eta Mutrikuko Hitza, Oarsoaldeko Hitza and Urola Kostako Hitza.

In this paper, we take representation learning in Basque as a case study, and explore tailored crawling (i.e., manually identifying and scraping websites with high-quality content) as an alternative to filtering CommonCrawl. We introduce EusCrawl, a new corpus for Basque comprising 12.5M documents from 33 websites with Creative Commons content. EusCrawl is similar in size to the Basque portion of CC100 and mC4, but it has substantially less issues and a higher perceived quality according to our blind audit with native annotators. However, we find that this improvement does not carry over to downstream NLU tasks, as masked language models pre-trained on either corpora obtain sim-

ilar results on 5 benchmarks. Our results suggests that data quantity and domain coverage play a more important role, prompting for methods to exploit more diverse sources of data in low-resource languages.

This paper makes the following contributions: (i) we release EusCrawl, a high-quality corpus for Basque comprising 12.5M documents and 423M tokens;[1] (ii) we manually assess the quality of EusCrawl in comparison with mC4 and CC100, finding that it has substantially less issues and a higher perceived quality according to native annotators; (iii) we compare masked language models pre-trained on EusCrawl, mC4, CC100 and Wikipedia[2] on 5 NLU tasks, finding that they all perform similarly with the exception of Wikipedia; and (iv) we obtain state-of-the-art results on several NLU benchmarks in Basque, outperforming prior work that relied on non-public corpora.

2 Experimental setup

We next detail the corpora compared in our experiments (§2.1), and the qualitative and downstream evaluation settings (§2.2 and §2.3).

2.1 Corpora

We compare 4 Basque corpora in our experiments: mC4, CC100, Wikipedia and EusCrawl. Table 1 summarizes their details. **mC4**[3] and **CC100**[4] are, to the best of our knowledge, the two largest public corpora for Basque. They were introduced to train mT5 (Xue *et al.*, 2021) and XLM-R (Conneau *et al.*, 2020), respectively, and were built by filtering CommonCrawl. **Wikipedia** has been a popular source for multilingual data (Pires *et al.*, 2019; Conneau and Lample, 2019; Artetxe *et al.*, 2020). We extract text from a Wikipedia dump using the WikiExtractor tool.[5] **EusCrawl** is a new corpus we introduce. Instead of filtering CommonCrawl, we do tailored crawling on 33 websites with high-quality content in Basque, mostly on the news domain. We build ad-hoc scrapers to extract text from these websites, resulting in higher coverage[6]and cleaner text compared to general purpose approaches. We only use content with a Creative Commons license. Table 2 summarizes all the sources we use.

2.2 Qualitative evaluation

We manually audit the quality of EusCrawl in comparison with mC4 and CC100 by randomly sampling 100 documents from each corpus (a total of 300 documents), and asking native annotators to assess their quality.[7] We ensure that the evaluation is blind by showing the documents in a random order and not revealing what corpus they were sampled from. For each document, we ask the annotators to assess if the document has any problem in each of the following categories: **langID** (the document is not in Basque), **language variety** (the document is not written in standard and correct Basque), **coherence** (the document has gaps and/or some portions are not connected), **noise** (the document is not clean) and **content** (the document seems to have been generated automatically and/or has no meat). In addition, we ask annotators to classify each document according to its **perceived quality** as high-quality (the document does not have quality issues and the annotator thinks that it would be good to have it in the corpus), medium-quality (the document has some minor issues and the annotator is unsure if it would be good to have it in the corpus), or low-quality (the document has major issues

[1] https://www.ixa.eus/euscrawl/. Meta AI was not involved in the collection and distribution of the corpus.

[2] Models available at https://dl.fbaipublicfiles.com/euscrawl/roberta-eus-{euscrawl|mc4|cc100|wikipedia}-{base|large}.tar.gz.

[3] We use the version released by AllenAI at https://github.com/allenai/allennlp/discussions/5265

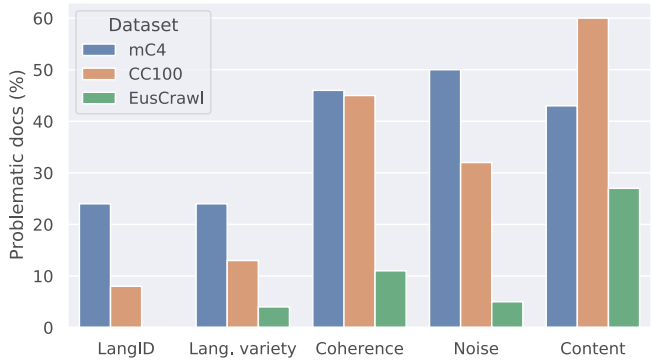
[4] We use the version from https://data.statmt.org/cc-100/

[5] https://github.com/attardi/wikiextractor

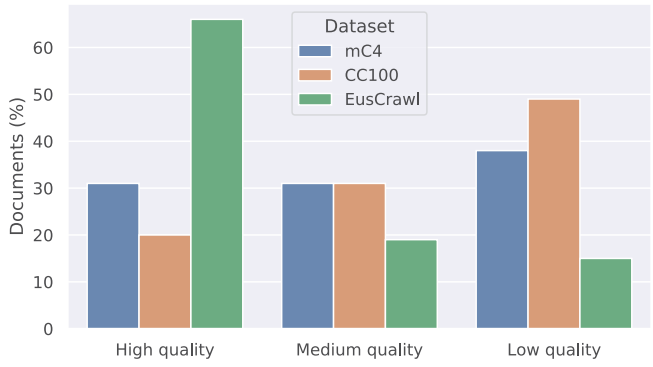
[6] While one may expect the websites we crawl to be covered by mC4 and CC100, a large fraction of this content is missing in them. This is both because CommonCrawl is far from being a complete dump of the Internet, and the filtering applied by CC100 and mC4 is noisy, removing valid content.

[7] So as to control for the variance across annotators, we asked two additional native speakers to evaluate a random subset of 100 documents. The main findings were consistent across all the 3 annotations, so we omit results for brevity.

and the annotator thinks that it would be better not to have it in the corpus). Refer to Appendix A for the complete instructions given to annotators.



(a) Issues



(b) Overall quality

Figure 1. Data audit results. EusCrawl has a much higher quality than mC4 and CC100. See §2.2 for more details

2.3. *Downstream evaluation*

In addition to the qualitative evaluation, we pre-train RoBERTa models (Liu *et al.*, 2019) on each corpus, and evaluate fine-tuning them on the following **NLU benchmarks**: topic classification on BHTC (Agerri *et al.*, 2020), sentiment classification on Behagune (Agerri *et al.*, 2020), stance detection on VaxxStance (Agerri *et al.*, 2021), Named Entity Recognition (NER) on EIEC (Alegria *et al.*, 2006), and extractive conversational Question Answering (QA) on Elkarrizketak (Otegi *et al.*, 2020). We provide additional details on these datasets in Appendix B.

We **pre-train** each model for 125k steps with a batch size of 2048 and a sequence length of 512, using the same hyperparameters as Liu *et al.* (2019). We train RoBERTa-base models for our main comparison using a learning rate of 7e-4, and further train a RoBERTa-large model on EusCrawl with a learning rate of 4e-4 to understand the effect of scaling. In all cases, we use the final checkpoint without early stopping. We use SentencePiece (Kudo and Richardson, 2018) for tokenization, using a 50k vocabulary learned in each separate corpus.

For **fine-tuning**, we use the same hyperparameters as Agerri *et al.* (2020). For topic classification, sentiment classification and stance detection, we use a batch size of 16, a learning rate of 2e-5 with linear decay and a warmup of 6%, and train the model for 10 epochs. For NER and QA, we use a batch size of 32, a constant learning rate of 5e-5, and train for 4 epochs. We did not perform any hyperparameter tuning or model selection, and report results on the test set. The development sets, when available, were not used.

3. Results

3.1. *Qualitative evaluation*

As shown in Figure 1, EusCrawl has the best quality by a large margin in all the axes that we consider. mC4 has a slightly higher perceived quality and less content-related issues than CC100, but more problematic documents in the other categories.

More concretely, we find that both mC4 and CC100 have a high proportion of documents with coherence, noise and content-related issues. In addition, mC4 has a significant number of langID and language variety problems. In contrast, EusCrawl has minimal issues in all categories but content, where it still does substantially better than mC4 and CC100. Taking a closer look, we find that most of these content-related issues in EusCrawl correspond to short, template-based Wikipedia articles (e.g., *Placosoma is a a genus of lizards in the family Gymnophthalmidae. They live in Brazil.*[8]), which should be easy to filter in future iterations. Finally, we find

[8] Original text in Basque: Placosoma Gymnophthalmidae familiako narrasti genero bat da. Brasilen bizi dira.

that the overall quality of EusCrawl documents is also much better according to native annotators, with approximately two thirds of the documents being annotated as high-quality, compared to less than one third for both mC4 and CC100.

All in all, our qualitative evaluation provides further evidence that multilingual corpora derived from CommonCrawl have major quality issues, and shows that tailored crawling can be an effective alternative to obtain high-quality data.

Table 3
Downstream results

		Topic class.	Sentiment	Stance det.	NER	QA	Avg
Prior best	Agerri *et al.* (2020)	76.8	78.1	–	87.1	–	–
	Otegi *et al.* (2020)	–	–	–	–	35.0	–
	Lai *et al.* (2021)	–	–	57.3†	–	–	–
RoBERTa-base	mC4	75.3 ±0.7	**80.4** ±1.5	59.1 ±5.2	86.0 ±1.0	35.2 ±1.8	67.2
	CC100	76.2 ±0.4	78.8 ±1.2	**63.4** ±3.5	85.2 ±1.2	35.8 ±1.1	67.9
	Wikipedia	70.0 ±0.8	72.4 ±2.3	53.2 ±4.6	71.6 ±13.1	27.4 ±0.2	58.9
	EusCrawl	76.2 ±0.6	77.7 ±1.4	57.4 ±4.7	86.8 ±0.6	34.6 ±1.8	66.5
RoBERTa-large	EusCrawl	**77.6** ±0.5	78.8 ±0.9	62.9 ±2.3	**87.2** ±0.4	**38.3** ±1.3	**69.0**

We report average F1 and standard deviation across 5 runs (micro F1 in all tasks except stance detection, where we report macro F1 of the *favor* and *against* classes following common practice). †Best result among systems that rely exclusively on textual data.

3.2. Downstream tasks

We report our downstream results in Table 3.

In contrast with the qualitative evaluation, we find that there is not a clear winner among mC4, CC100 and EusCrawl. In fact, when looking at RoBERTa-base results, we find that mC4 does the best on sentiment classification, CC100 does the best on stance detection and QA, and EusCrawl does the best on NER. Wikipedia lags behind them all by a large margin. It is worth noting that the variance is high in certain tasks, which we attribute to the small size of the test sets and their unbalanced nature, but the general trends are consistent.

These results suggest that corpus quality issues in low-resource languages do not have a a major impact on NLU performance. Instead, we find evidence that it is the size and domain of the training corpus that is more important. This would explain why Wikipedia obtains the worst results, as it is substantially smaller than the other corpora and restricted to a narrow domain. Similarly, this is also consistent with EusCrawl performing worse than mC4 and CC100 on sentiment analysis and stance detection, as the domain of these benchmarks (tweets) is different from the domain of EusCrawl (primarily news, see Table 2), while CommonCrawl-derived corpora are presumably more diverse.

Finally, we find that scaling to RoBERTa-large brings consistent improvements in all tasks. Thanks to this, we are able to outperform the best published results in all the 5 benchmarks. Note that we achieve this pre-training exclusively on Creative Commons data that we release publicly, while prior work relied on private datasets.

4. Conclusions

Taking Basque as a case study, our work gives further evidence that CommonCrawl-derived corpora have major quality issues in low-resource languages. At the same time, we show that ad-hoc crawling websites with high-quality content can be an effective alternative to collect data in such languages. Our resulting corpus EusCrawl has a higher quality than mC4 and CC100 according to our manual data audit, while being similar in size. Nevertheless, this improvement in quality does not carry over to downstream performance on NLU tasks, where we find evidence that data quantity and domain coverage are more important factors.

Our work leaves important lessons for future efforts on low-resource languages. First of all, we find that, even if CommonCrawl derived multilingual corpora do have major quality issues as raised by prior work (Kreutzer *et al.*, 2021), these issues do not have a significant impact in NLU tasks. This suggests that investing on bigger and more diverse datasets might be more fruitful than addressing such quality issues in low-resource settings. Given that the amount of written text in such languages is ultimately limited, we believe that developing effective cross-lingual transfer methods to exploit multilingual data is a promising future direction. Having said that, it should be noted that our study is limited to NLU tasks in a single language. It is possible that data quality plays a more important role in generation tasks, which we leave for future work to study. In addition, we think that it would be valuable to conduct similar studies in other languages to corroborate our findings.

Finally, we note that prior work on Basque NLP has often relied on private resources (Agerri *et al.*, 2020). Our work sets a new state-of-the-art on a diverse set of NLU benchmarks, and it does so using public data alone. By releasing our corpus, we hope to facilitate future work in Basque NLP, and encourage open and reproducible science using public resources.

5. Limitations

Our evaluation focuses on NLU tasks, and it is possible that data quality plays a different role in generation tasks. We note, however, that generation quality is harder to evaluate through automatic metrics, which is why we decided to focus on NLU tasks. Moreover, the corpora that we compare differ on various aspects other than the data quality (e.g., the domain), and it is hard to isolate the effect of quality from the rest. In any case, we believe that our main claim still holds, in that data quality has a minor impact relative to such other factors. Finally, our work builds on EusCrawl —a new high-quality corpus that we introduce for Basque— and our analysis is thus limited to this language. It would be interesting to collect high-quality corpora for other low-resource languages, and conduct a similar comparison to corroborate that our findings also apply more broadly.

6. Acknowledgments

We would like to thank Nikolas Vicuña for his help on expanding EusCrawl, Naman Goyal for his advice on pre-training RoBERTa, and Arantza Rico and Paloma Rodriguez-Miñambres for their help with the manual data audit. We are also grateful to all Basque media and content creators that share their work under a Creative Commons license, making a resource like EusCrawl possible.

Itziar Aldabe, Rodrigo Agerri, Olatz Perez-de-Viñaspre and Aitor Soroa were supported by the Basque Government (excellence research group IT1343-19 and DeepText project KK-2020/00088), and by the projects: (i) Deep-Knowledge (PID2021-127777OB-C21) funded by MCIN/AEI/10.13039/501100011033 and FEDER Una manera de hacer Europa; (ii) Disargue (TED2021-130810B-C21), MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR; (iii) DeepR3 (TED2021-130295B-C31) funded by MCIN/AEI/10.13039/501100011033 and EU NextGeneration programme EU/PRTR. Rodrigo Agerri's work is also supported by the RYC-2017-23647 fellowship (MCIN/AEI /10.13039/501100011033 y por El FSE invierte en tu futuro).

7. References

Rodrigo Agerri, Roberto Centeno, María Espinosa, Joseba Fernandez de Landa, and Álvaro Rodrigo Yuste. 2021. Vaxxstance@iberlef 2021: Overview of the task on going beyond text in cross-lingual stance detection.

Rodrigo Agerri, Iñaki San Vicente, Jon Ander Campos, Ander Barrena, Xabier Saralegi, Aitor Soroa, and Eneko Agirre. 2020. Give your text representation models some love: the case for basque. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 4781-4788, Marseille, France. European Language Resources Association.

Iñaki Alegria, Olatz Arregi, Nerea Ezeiza, and Izaskun Fernández. 2006. Lessons from the development of a named entity recognizer for Basque. *Procesamiento del Lenguaje Natural*, 36: 25-37.

Mikel Artetxe, Sebastian Ruder, and Dani Yogatama. 2020. On the cross-lingual transferability of mono-lingual representations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4623-4637, Online. Association for Computational Linguistics.

Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri Chatterji, Annie Chen, Kathleen Creel, Jared Quincy Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, Omar Khattab, Pang Wei Koh, Mark Krass, Ranjay Krishna, Rohith Kuditipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Ben Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, Julian Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Rob Reich, Hongyu Ren, Frieda Rong, Yusuf Roohani, Camilo Ruiz, Jack Ryan, Christopher Ré, Dorsa Sadigh, Shiori Sagawa, Keshav Santhanam, Andy Shih, Krishnan Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. 2021. On the opportunities and risks of foundation models.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440-8451, Online. Association for Computational Linguistics.

Alexis Conneau and Guillaume Lample. 2019. Cross-lingual language model pretraining. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. 2020. Scaling laws for neural language models.

Julia Kreutzer, Isaac Caswell, Lisa Wang, Ahsan Wahab, Daan van Esch, Nasanbayar Ulzii-Orshikh, Allahsera Tapo, Nishant Subramani, Artem Sokolov, Claytone Sikasote, Monang Setyawan, Supheakmungkol Sarin, Sokhar Samb, Benoît Sagot, Clara Rivera, Annette Rios, Isabel Papadimitriou, Salomey Osei, Pedro Ortiz Suárez, Iroro Orife, Kelechi Ogueji, Andre Niyongabo Rubungo, Toan Q. Nguyen, Mathias Müller, André Müller, Shamsuddeen Hassan Muhammad, Nanda Muhammad, Ayanda Mnyakeni, Jamshidbek Mirzakhalov, Tapiwanashe Matangira, Colin Leong, Nze Lawson, Sneha Kudugunta, Yacine Jernite, Mathias Jenny, Orhan Firat, Bonaventure F. P. Dossou, Sakhile Dlamini, Nisansa de Silva, Sakine Çabuk Ballı, Stella Biderman, Alessia Battisti, Ahmed Baruwa, Ankur Bapna, Pallavi Baljekar, Israel Abebe Azime, Ayodele Awokoya, Duygu Ataman, Orevaoghene Ahia, Oghenefego Ahia, Sweta Agrawal, and Mofetoluwa Adeyemi. 2021. Quality at a glance: An audit of web-crawled multilingual datasets.

Taku Kudo and John Richardson. 2018. SentencePiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 66-71, Brussels, Belgium. Association for Computational Linguistics.

Mirko Lai, Alessandra Teresa Cignarella, Livio Finos, and Andrea Sciandra. 2021. Wordup! at vaxxstance 2021: Combining contextual information with textual and dependency-based syntactic features for stance detection. In *IberLEF@SEPLN*.

Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O'Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, and Xian Li. 2021. Few-shot learning with multilingual language models.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach.

Arantxa Otegi, Aitor Agirre, Jon Ander Campos, Aitor Soroa, and Eneko Agirre. 2020. Conversational question answering in low resource scenarios: A dataset and case study for basque. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 436-442, Marseille, France. European Language Resources Association.

Telmo Pires, Eva Schlinger, and Dan Garrette. 2019. How multilingual is multilingual BERT? In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4996–5001, Florence, Italy. Association for Computational Linguistics.

Jack W. Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, Eliza Rutherford, Tom Hennigan, Jacob Menick, Albin Cassirer, Richard Powell, George van den Driessche, Lisa Anne Hendricks, Maribeth Rauh, Po-Sen Huang, Amelia Glaese, Johannes Welbl, Sumanth Dathathri, Saffron Huang, Jonathan Uesato, John Mellor, Irina Higgins, Antonia Creswell, Nat McAleese, Amy Wu, Erich Elsen, Siddhant Jayakumar, Elena Buchatskaya, David Budden, Esme Sutherland, Karen Simonyan, Michela Paganini, Laurent Sifre, Lena Martens, Xiang Lorraine Li, Adhiguna Kuncoro, Aida Nematzadeh, Elena Gribovskaya, Domenic Donato, Angeliki Lazaridou, Arthur Mensch, Jean-Baptiste Lespiau, Maria Tsimpoukelli, Nikolai Grigorev, Doug Fritz, Thibault Sottiaux, Mantas Pajarskas, Toby Pohlen, Zhitao Gong, Daniel Toyama, Cyprien de Masson d'Autume, Yujia Li, Tayfun Terzi, Vladimir Mikulik, Igor Babuschkin, Aidan Clark, Diego de Las Casas, Aurelia Guy, Chris Jones, James Bradbury, Matthew Johnson, Blake Hechtman, Laura Weidinger, Iason Gabriel, William Isaac, Ed Lockhart, Simon Osindero, Laura Rimell, Chris Dyer, Oriol Vinyals, Kareem Ayoub, Jeff Stanway, Lorrayne Bennett, Demis Hassabis, Koray Kavukcuoglu, and Geoffrey Irving. 2022. Scaling language models: Methods, analysis & insights from training gopher.

Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. mT5: A massively multilingual pre-trained text-to-text transformer. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 483-498, Online. Association for Computational Linguistics.

iStock

Nola jarraitu birus baten hedapenari pribatutasun-arazorik sortu gabe

COVID-19aren pandemiaren eboluzioa aztertzeko, ezinbestekoa da birusaren hedapenari jarraitzea. Baina jarraitze-metodoek pribatutasun-arazo larriak eragin ditzakete, adibidez, jarraipena irizpide geografikoen arabera egiten bada. Horregatik, matematika erabilita, beste irizpide batzuen araberako metodoak garatu dira. Garrantzitsua da, pandemia batean ere, pertsonen pribatutasunaren alde egiten duten irizpideak ezartzea gaixotasun baten hedapenaren berri izan nahi denean.

Osasunaren Mundu Erakundeak pandemia deklaratu zuen 2020ko martxoan, eta birusak 117.700 milioi kasu baino gehiago eta 2.600 milioi heriotza eragin zituen 2021eko martxorako. Gaixotasunak edozein pertsonari eragin diezaioke; hala ere, badira arrisku handiko pertsona-talde batzuk, batez ere pertsona adinduak eta beste gaixotasun batzuk dituztenak. Testuinguru horretan, mundu osoko gobernuek zenbait erabaki hartu zituzten pandemia gehiago zabaltzea eragozteko: jendea konfinatzea, urruntze soziala eta abar. Neurri haietako bakoitzak ondorio ekonomiko batzuk zituen.

Oro har, pandemiaren eboluzioari irizpide geografikoen arabera jarraitu zaio, gune bateko populazioan bereizketarik egin gabe, eta gaixotutakoen jarraipena egiteko erraminta digitalek pribatutasun-kezka handiak eragin dituzte. Telefono mugikorretako aplikazioen bidez, eta telefono bakoitzaren Bluetooth seinalea baliatuta, pertsonen arteko gertutasuna kontrolatu izan da, infekzio-bideak atzemateko. Nolanahi ere, aplikazio horiek pertsonen informazio osoa jartzen diete eskura agintariei, eta pribatutasuna galdu egiten da.

Talde ez-geografikoak

Arazo horren aurrean, algoritmo berezi bat erabiltzea proposatu dute zenbait adituk. Aspalditik erabiltzen dira datu pertsonalak babesten dituzten metodo informatikoak, eta algoritmo horrek, hain zuzen ere, metodo horietako batzuk baliatzen ditu. Haien bitartez, pertsonak taldekatzen ditu, haien kontaktuak zein diren eta nolako arrisku-maila duten aztertuta. Horrela, datu pertsonalak babesten dituzten taldeak sortzen ditu. Talde bakoitzean adostasun-algoritmo bat erabiltzean, gizabanakoek taldearen egoera epidemiologikoari buruzko informazioa izan dezakete, eta, ondorioz, urruntze sozialeko neurriak egokitu. Ikuspegi espezifiko horrek bermatzen du arrisku handiko taldeetan soilik hartzea neurri zorrotzagoak, eta, hala, murrizketa geografiko zabalekin lotutako eragin ekonomikoa arintzen da.

«Algoritmo honek kontaktuen eta arrisku-mailaren arabera taldekatzen ditu pertsonak, baina pribatutasun pertsonala puskatu gabe»

Harrigarria bada ere, algoritmoak pribatutasun indibidualari eusten dio, eta erakunde zentralik gabe jarduten du; izan ere, pertsona bakoitzak bere taldeko afiliazioaz baino ez du izan behar kontziente, eta ez taldeko kide zehatzez. Algoritmoaren moldagarritasuna ezinbestekoa da, taldeak etengabe doitzen baititu gizarte-harremanen aldaketei eta arrisku-mailei erantzuteko, txertaketaren aurrerapena barne. Konplexutasun konputazionalaren analisiak algoritmoaren eraginkortasuna berresten du, haren baliabide-eskaerak populazioaren tamainarekin hazten baitira.

Horrelako metodoak erabilita, agintariek ez dute datu pertsonalik jaso behar taldeei segimendua egiteko. Beste pandemia bat baletor, horrelako tresnak erabili beharko lituzkete gobernuek, eraginkortasun handiko jarraipena egiteko, pribatutasuna puskatu gabe.

InpAkta 00 (2023) 308-320

https://doi.org/10.1109/ACCESS.2021.3131777

Distributed clustering algorithm for adaptive pandemic control

Xabier Insausti*, Marta Zárraga-Rodríguez, Carolina Nolasco-Ferencikova and Jesús Gutiérrez-Gutiérrez

Tecnun University of Navarra

ABSTRACT: The COVID-19 pandemic has had severe consequences on the global economy, mainly due to indiscriminate geographical lockdowns. Moreover, the digital tracking tools developed to survey the spread of the virus have generated serious privacy concerns. In this paper, we present an algorithm that adaptively groups individuals according to their social contacts and their risk level of severe illness from COVID-19, instead of geographical criteria. The algorithm is fully distributed and therefore, individuals do not know any information about the group they belong to. Thus, we present a distributed clustering algorithm for adaptive pandemic control.[1]

[1] This work was supported in part by the Spanish Ministry of Science and Innovation through the ADELE project (PID2019-104958RB-C44).

1 Introduction

COVID-19 [1] is a disease caused by the new coronavirus SARS-CoV-2. It was declared a pandemic by the World Health Organization (WHO) in March 2020. First cases were reported in Wuhan, People's Republic of China, to the WHO on December 31st 2019. Since then, 117.7 billion cases have been reported, with more than 2.6 billion deaths, as of March 10th, 2021 [2]. Those at a higher risk of severe illness from COVID-19 include those aged 60 or over, or with underlying medical problems such as diabetes, cancer, or high-blood pressure. Nevertheless, this highly infectious disease can affect anyone, and can become deadly at any age. Personal health precautions are strongly advised, mainly wearing a mask, physical distancing and handwashing [1].

* **Corresponding author / Harremanetan jartzeko:** Xabier Insausti. Tecnun, University of Navarra, San Sebastián. - https://orcid.org/0000-0001-9628-0681. **Contributed by** Xabier Insausti Sarasola **editoreak egindako ekarpena**.

How to cite / Nola aipatu: X. Insausti, *et al.* Nolasco-Ferencikova and J. Gutiérrez-Gutiérrez, «Distributed Clustering Algorithm for Adaptive Pandemic Control», in *IEEE Access*, 2021, vol. 9, 160688-160696 (https://doi.org/10.1109/ACCESS.2021.3131777)

In response to the pandemic, governments all over the world have implemented non-pharmaceutical measures in order to stop the spread of the virus, or *flatten the curve*. Social distancing interventions, such as isolation and quarantine of infected patients and their contacts, external and internal border restrictions, workplace distancing, closure of schools, and complete quarantine or lockdown have been the most common [3, 4]. FluTE, a stochastic influenza pandemic simulation model [5], was used to assess the potential effect of different social distancing interventions using Singapore

as a study case [6], since it was among the first to report infections. The model predicted quarantine or lockdown to be the most effective measures, particularly combined with school closures and workplace distancing. In fact, Singapore successfully implemented these measures, preventing community spread [6]. It is important to point out that these measures are targeted geographically [7]. This geographical approach affects large population groups, regardless of their economic sector or activity. Therefore, these measures have severe consequences on the regional, national and global economy: they pose a risk of reduced income or even job loss, affecting the most dis- advantaged populations [8], and results show an average 2.5-3% global GDP drop per month of complete lockdown [9]. This shows that, despite lockdown and quarantine being the most effective measures, a different non-geographical approach should be taken in order to overcome the aforementioned negative impacts. Furthermore, these measures are most efficient when applied to individuals that belong to groups where transmission is most likely to occur [10]. Hence, individuals should be grouped according to their social contacts, which might not necessarily coincide with geographical areas. However, if the criteria are not geographical, it is more difficult for individuals to know which group they belong to. Furthermore, such groups may change with time and adaptive grouping strategies are needed.

Public health experts across institutions and countries have identified digital tracking measures as useful tools to survey and slow down the spread of the virus. Numerous technologies have been developed with this purpose, such as digital health certificates, which assign a color-coded COVID status to their users, physical surveillance initiatives [11], symptom checkers, or flow modelling tools, which quantify and track people's movements in specified geographical regions [12]. These technologies, however, raise severe ethical concerns about putting user's privacy and security at risk. For instance, out of the 65 digital health certificate applications that are currently in operation globally, 82% are considered to have inadequate privacy policies [11].

One of the most common examples of digital tracking measures are proximity or contact tracing tools, mainly via mobile applications. In particular, studies have predicted them to be beneficial in mitigating the spread of the virus, specifically during the de-escalation of physical distancing [13]. There are over 120 contact tracing applications currently available in over 70 countries [11]. These contact tracing tools gather data from their users, such as their location, their health records or contact information. This has raised ethical concerns surrounding the privacy of users and their data.

For instance, one of the earlier contact tracing tools developed was Singapore's TraceTogether [14], a mobile application which operates via Bluetooth connection. Nearby phones, with Bluetooth and TraceTogether open in the background, exchange tokens, which are stored encrypted in each phone and in a central server [15]. If a user tests positive for COVID-19, contact tracers can easily use the tokens to identify those at high risk of infection. TraceTogether does not gather more than the necessary information, only the users' contact/mobile number, identification details and random ID. The tokens sent via Bluetooth are time-varying random strings, and this way, privacy between users is kept. However, when a user is infected, the government can retrieve all mobile numbers of the individuals the infected user has been in contact with [15]. Having this centralized approach leaves no privacy for users from authorities.

For overcoming the privacy concerns of a centralized approach, in an unprecedented joint effort Apple and Google developed a contact tracing plat-

form based on Bluetooth [16]. Specifically, they developed an application programming interface (API) that allows interoperability between Android and IOS. This API requires contact tracing applications to take on a decentralized approach. The contact matching analysis is performed at a local level, which also protects users' privacy, maintaining their anonymity. Over 37% of contact tracing applications now use Apple and Google's API [11].

In this paper, we propose a distributed algorithm that adaptively groups individuals (i.e., creates clusters) according to their social contacts and their risk level of severe illness from COVID-19. This will be modelled as a doubly-weighted undirected graph. Moreover, by combining our algorithm with a distributed consensus algorithm, each individual can know the epidemiological situation of the group they belong to and can take the social distancing measures that correspond to the epidemiological situation of their group.

There exist many algorithms to create clusters and, in particular, many works about privacy- preserving clustering have been conducted (see, e.g., [17–21]). These works are based on statistical or cryptography techniques to protect data. Our algorithm can use some of the abovementioned techniques for becoming privacy-preserving between nearby users, but since it is fully distributed individuals do not share any information about the cluster they belong to even if no cryptographic methods are used. Therefore, privacy from authorities is kept. That is, only the individuals themselves know which group (cluster) they belong to without having knowledge of the rest of the members of the group.

In the literature, many works deal with distributed clustering of data using a wide variety of techniques and applying the results to different fields (see, e.g., [22-29]). In this paper, we focus on spectral clustering techniques because they are easy to implement and have been shown to be more effective in finding clusters than some traditional algorithms such as k-means [30]. Among the previously cited works, [27-29] present a similar approach to the one considered in this paper. Specifically, in [27] the authors propose a distributed spectral clustering algorithm but they do not consider weights neither in the nodes nor in the edges. In [28], the authors propose a distributed spectral clustering algorithm but they only consider an edge-weighted graph. Finally, in [29] a spectral clustering for doubly-weighted graphs is proposed but, unlike here, the algorithm is not distributed.

The remainder of this paper is organized as follows. Section 2 states preliminary considerations regarding distributed computation and spectral clustering. Section 3 presents the distributed clustering algorithm for adaptive pandemic control, its convergence speed, and its computational complexity. Finally, two illustrative examples and some conclusions are given in Sections 4 and 5, respectively.

2. Preliminaries

2.1. *Distributed computation using a linear iterative algorithm*

Consider a network composed of n nodes, where each node represents the mobile phone (or similar) device of one person. The entire population and the interactions among them can be viewed as an undirected graph $G = (\mathrm{V}, \mathcal{E})$ with no loops, where $\mathrm{V} = \{1, 2, \ldots, n\}$ is the set of nodes (vertices) and $\mathcal{E}$ is the set of edges. If two nodes $i, j \in \mathrm{V}$ interact between them, then $\{i, j\} \in \mathcal{E}$. We say that these nodes are connected, and can therefore interchange information. Conversely, if $\{\mathrm{i}, \mathrm{j}\} \notin \mathcal{E}$, this means that nodes $i, j \in \mathrm{V}$ are not connected and cannot interchange information.

We assume that each node $i \in \mathrm{V}$ has an initial value $x_i(0) \in \mathbb{R}$, where $\mathbb{R}$ denotes the set of real numbers. In distributed computation each node computes its target value by interchanging information with its neighbouring nodes. The approach that will be considered here for distributed computation is to use a linear iterative algorithm of the form

$$x_i(t+1) = w_{i,i}x_i(t) + \sum_{j\in\mathcal{V}:\{i,\mathrm{j}\}\in\mathcal{E}} w_{i,j}x_j(t), \qquad i \in \mathcal{V}, \tag{1}$$

where $w_{i,j} \in \mathbb{R}$ and time $t \in \{0, 1, 2, \ldots\}$ is assumed to be discrete (see [31]). Let $x(t) = (x_1(t), x_2(\mathrm{t}), \ldots, x_n(t))^\top$ be the column vector with the values of the nodes

at time instant t, where $\top$ denotes transpose. The linear iterative algorithm (1) can then be written as

$$\boldsymbol{x}(t+1) = W\,\boldsymbol{x}(t) = W^{t+1}\,\boldsymbol{x}(0), \tag{2}$$

where W is the $n \times n$ matrix defined as

$$[W]_{i,j} = \begin{cases} 0 & \text{if } \ i \neq j \text{ and } \{i,j\} \notin \mathcal{E}, \\ w_{i,j} & \text{otherwise,} \end{cases} \tag{3}$$

for $i, j \in \mathrm{V}$.

2.2. *Spectral clustering*

Clustering a graph consists in separating the nodes of the graph into disjoint groups (clusters). There exist many algorithms for graph clustering. The approach that will be considered here is the so-called *spectral clustering* (see, e.g., [32–34]). Spectral clustering is based on the information provided by the eigenvectors of the Laplacian matrix of the graph [35], mainly by an eigenvector corresponding to the smallest nonzero eigenvalue of such matrix, known as *Fiedler vector* [36].

In this paper G is assumed to be a doubly-weighted graph, that is, a graph with weights both in the nodes and in the edges. We denote with $\delta_i > 0$ the weight of node i, for $i \in \mathrm{V}$, and whenever $\{\mathrm{i}, \mathrm{j}\} \in \mathcal{E}$ we denote with $\sigma_{i,j} > 0$ the weight of such edge.

In [29, Lemma 1], in the context of doubly-weighted graphs, the notion of weighted Laplacian matrix was presented. The weighted Laplacian matrix of the graph is the $n \times n$ matrix given by

$$L = \Lambda^{-\frac{1}{2}}(D - \Sigma)\Lambda^{-\frac{1}{2}}, \tag{4}$$

where Λ^{-2} is the $n \times n$ diagonal matrix with $[\Lambda^{-\frac{1}{2}}]_{i,i} = \dfrac{1}{\sqrt{\delta_i}}$,

$$[\Sigma]_{i,j} = \begin{cases} \sigma_{i,j} & \text{if } \ \{i,j\} \in \mathcal{E}, \\ 0 & \text{if } \ \{i,j\} \notin \mathcal{E}, \end{cases}$$

and D is the $n \times n$ diagonal matrix with $[D]_{i,i} = \sum_{j=1}^{n} [\Sigma]_{i,j}$.

From [37, Theorem 5.1], L is positive semidefinite. Let $L = U\,\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda n)U^{-1}$ be an eigenvalue decomposition of L, where the eigenvalues are arranged in non-decreasing order and the eigenvector matrix $U = U = [\boldsymbol{u}_1 | \boldsymbol{u}_2 | \ldots | \boldsymbol{u}_n]$ is real and orthogonal. Assume that G has k components. Then, $\lambda_1 = \ldots = \lambda_k = 0$. In [37, Section 5.1] it is shown that $[\boldsymbol{u}_{k+1}]_i$ indicates which cluster the node i belongs to.

3. Distributed clustering algorithm for adaptive pandemic control

3.1. *Proposed algorithm*

Consider a set of n individuals that interact in a certain geographical region. The entire population and the interactions among them will be modelled with a doubly-weighted undirected graph G with no loops. The node i of the graph represent the i-th individual and the weight of the node i, δ_i, represents the individual's risk level of severe illness from COVID-19. The edge $\{i, j\}$ of the graph represents that there exists an interaction between individuals i and j, and the weight of the edge, $\sigma_{i,j}$, represents the time frame of the social contact between them.

In this section we present an algorithm that adaptively groups individuals according to their social contacts and their risk level of severe illness from COVID-19, that is, we present an algorithm for clustering the doubly-weighted graph G. Since the goal is to keep privacy from authorities, the algorithm presented here is fully distributed. Specifically, it computes the eigenvector $\boldsymbol{u}_{k+1}$ of the Laplacian matrix L of the graph G in a distributed way (see Algorithm 1).

The rest of this section is devoted to proving that $\boldsymbol{u}k{+}1$ can be computed in a distributed way (Theorem 1). Theorem 1 directly provides the steps of Algorithm 1.

Theorem 1 *Consider a doubly-weighted undirected graph G with no loops, n nodes, and k components. Let the Laplacian matrix L of the graph G be as in* (4) *with* $\lambda k{+}1 < \lambda k{+}2$ *. Then, for almost every column vector* $\boldsymbol{u}(0)$,

$$\lim_{t\to\infty} \big[\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\big]_i \left(\frac{\big[\boldsymbol{x}(t-1) - \boldsymbol{x}(t-2)\big]_i}{\big[\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\big]_i} \right)^{t-1} = C\big[\boldsymbol{u}_{k+1}\big]_i \tag{5}$$

for all $i \in \mathrm{V}$, *where*

$$\boldsymbol{x}(t+1) = \left(I_n - \frac{1}{\lambda_n} L\right) \boldsymbol{x}(t) \qquad \forall t \in \{0,1,2,\ldots\}, \tag{6}$$

C is a non-zero constant, and In denotes the n × n identity matrix.

Proof: See Appendix 6.

In the rare case in which $\lambda_{k+1} = \lambda_{k+2}$ the Fiedler vector would not be unique, meaning that it might be any vector in a subspace of dimension larger than one. In this rare case, Algorithm 1 would still work because it would converge to one of such vectors.

Observe that the iterative equation (6) can be computed in a distributed way since it is of the form of (2), and $\left(I_n - \frac{1}{\lambda_n} L\right)$ satisfies (3). Therefore, from (5) each node $i \in \mathrm{V}$ can know the i-th entry of an eigenvector associated to λ_{k+1}. However, in order to compute (6) in a distributed way, each node needs to know λ_n. Lemma 1 shows that λ_n can also be computed in a distributed way.

Lemma 1 *Consider a doubly-weighted undirected graph G with no loops, n nodes, and k components. Let the Laplacian matrix L of the graph G be as in (4). Then, for almost every real n-dimensional column vector* $\boldsymbol{y}(0)$,

$$\lim_{t\to\infty} \frac{[\boldsymbol{y}(t)]_i}{[\boldsymbol{y}(t-1)]_i} = \lambda_n \qquad \forall i \in \mathcal{V}, \tag{7}$$

where

$$\boldsymbol{y}(t+1) = L\boldsymbol{y}(t) \qquad \forall t \in \{0,1,2,\ldots\}. \tag{8}$$

Proof: See [38, Section 5.8.1] or [39, Section 9.3].

Observe that the iterative equation (8) can be computed in a distributed way since it is of the form of (2), and L satisfies (3). Therefore, from (7) each node $i \in \mathrm{V}$ can know λ_n.

It should be mentioned that the distributed computation of u_{k+1} can be found in [28], but only for an edge-weighted graph, that is, for the particular case in which $\delta_i = 1$ for all $i \in \mathrm{V}$.

We finish this section by describing Algorithm 1. For ease of notation, we define

$$f(\boldsymbol{x}, t) := W^t \boldsymbol{x}(0),$$

which is the t-th iteration of (1) and can clearly be computed in a distributed way. As for Algorithm 1, we fix t_0 to be the number of iterations of (1) required for a desired precision. Table 1 describes Algorithm 1 and relates it with the theoretical aspects shown in this section. Observe that Algorithm 1 separates the nodes of the graph into two clusters. However, if the algorithm is used recursively within each cluster, we can separate the nodes of the original graph into as many clusters as desired.

Table 1
Explanation of Algorithm 1

Lines	Description
1-7	In (2), set W as L to compute (8)
10-12	Computation of λ_n according to Lemma 1
13-17	In (2), set W as $\left(I_n - \frac{1}{\lambda_n} L\right)$ to compute (6)
20-23	Computation of the i-th entry of an eigenvector associated to λ_2 according to
24-26	Assign node i to a cluster depending on the sign of $[u_{k+1}]i$

Algorithm 1
Distributed clustering algorithm for adaptive pandemic control

for all nodes $i \in \mathcal{V}$ **do**
  $s_i \leftarrow 0$
  for all nodes j connected to i **do**
    $w_{i,j} \leftarrow \frac{-\sigma_{i,j}}{\sqrt{\delta_i \delta_j}}$
    $s_i \leftarrow s_i + \sigma_{i,j}$
  end for
  $w_{i,i} \leftarrow \frac{s_i}{\delta_i}$
  $[\boldsymbol{y}(0)]_i \leftarrow \text{rand}()$ > An arbitrary value
end for
for all nodes $i \in \mathcal{V}$ **do**
  $\lambda_n \leftarrow \frac{[f(\boldsymbol{y}, t_0)]_i}{[f(\boldsymbol{y}, t_0 - 1)]_i}$
end for
for all nodes $i \in \mathcal{V}$ **do**
  for all nodes j connected to i **do**
    $w_{i,j} \leftarrow -\frac{w_{i,j}}{\lambda_n}$
  end for
  $w_{i,i} \leftarrow 1 - \frac{w_{i,i}}{\lambda_n}$
  $[\boldsymbol{x}(0)]_i \leftarrow \text{rand}()$ 1 > An arbitrary value
end for
for all nodes $i \in \mathcal{V}$ **do**
  $\beta_i \leftarrow [f(\boldsymbol{x}, t_0)]_i - [f(\boldsymbol{x}, t_0 - 1)]_i$
  $\gamma_i \leftarrow [f(\boldsymbol{x}, t_0 - 1)]_i - [f(\boldsymbol{x}, t_0 - 2)]_i$
  $[C\boldsymbol{u}_{k+1}]_i \leftarrow \beta_i \left(\frac{\gamma_i}{\beta_i}\right)^{t_0 - 1}$
  if $[C\boldsymbol{u}_{k+1}]_i > 0$ then node i belongs to cluster 1
  else node i belongs to cluster 2
  end if
end for

3.2. Convergence speed

In this subsection we study the convergence speed of the proposed algorithm. Specifically, we show that the convergence of the sequences considered in Theorem 1 and Lemma 1 is linear. We recall that the convergence of a sequence $a_0, a_1, a_2, \ldots$, which converges to ℓ, is said to be linear if the limit

$$\lim_{t\to\infty}\frac{|a_{t+1}-\ell|}{|a_t-\ell|}$$

is a nonzero constant (see [38, p. 224]).

The following theorem deals with the convergence speed of the sequence considered in Theorem 1.

Theorem 2 *Let x(t) be as in Theorem 1. Then, the convergence of the sequence*

$$\left[\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\right]_i\left(\frac{\left[\boldsymbol{x}(t-1)-\boldsymbol{x}(t-2)\right]_i}{\left[\boldsymbol{x}(t)-\boldsymbol{x}(t-1)\right]_i}\right)^{t-1}$$

is linear for all $i \in \mathrm{V}$.

Proof: See Appendix 7.

Since the convergence of the sequence considered in Lemma 1 is also linear (see [38, Section 5.8.1]), we conclude that the overall convergence of Algorithm 1 is linear.

3.3. Computational complexity

The computational bottleneck in spectral clustering is the computation of the eigenvectors of the Laplacian matrix. To speed up the computation of such eigenvectors, the power iteration method is usually used [40].

In this subsection we study the computational complexity of Algorithm 1 for each node. The computational complexity of Algorithm 1 is essentially determined by the complexity of running twice the power iteration method. In particular, the power iteration method is used to compute the largest eigenvalue of L (see line 11 of Algorithm 1) and to compute an eigenvector associated to the largest eigenvalue less than one of $I_n-\frac{1}{\lambda_n}L$ (see lines 21-22 of Algorithm 1). The power iteration method is computationally expensive for large matrices but L and $I_n-\frac{1}{\lambda_n}L$ are sparse matrices with only a few non-zero entries. This reduces the computational difficulties, as subsequently explained.

Let c_i be the number of contacts the i-th individual has. It is important to remark that c_i does not depend on n. Consequently, regardless of the value of n, the i-th row of L will have at most c_i+1 non-zero entries. Therefore, the computation of $[f(\boldsymbol{y}, t_0)]_i$ needed in line 11 requires no more than t0 (c_i+1) multiplications (see Equation (1)). Similarly, the computation of $[f(\boldsymbol{x}, t_0)]_i$ needed in lines 21-22 requires no more than t0 (c_i+1) multiplications.

Observe that t_0 controls the precision of the power iteration method and is usually not larger than 100 even for a very large n. Moreover, in [41] it is shown that even if n increases, t0 does not need to increase faster than $O(\log n)$ to keep the same precision. Consequently, in the worst case scenario, the computational complexity of Algorithm 1 is $O(\log n)$, which makes it suitable for a large n.

Finally, observe that regarding the memory usage of the algorithm, node i only needs to store c_i+1 values (the i-th row of L) and therefore the storage requirement of each node does not increase with n.

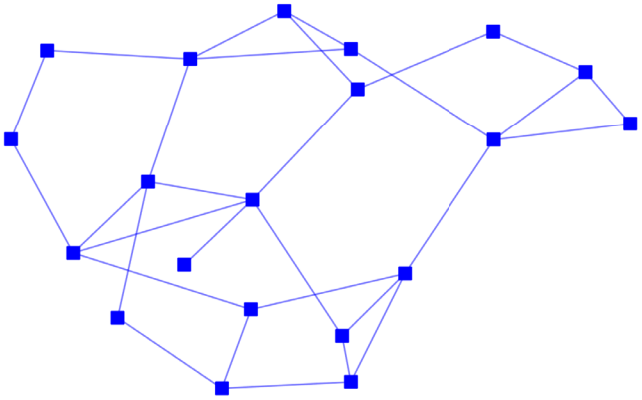
4. Illustrative examples

In this section we present two examples to illustrate how Algorithm 1 works.

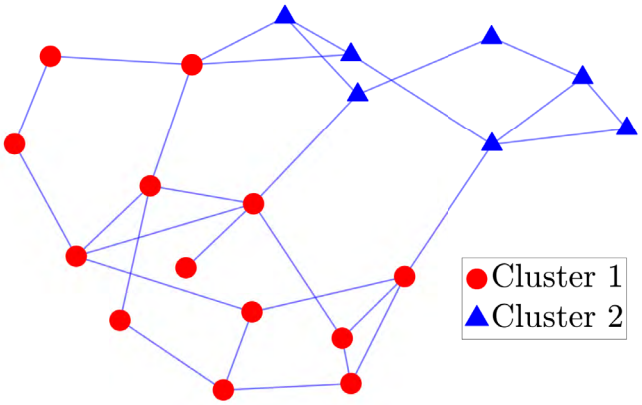
4.1. Example with randomly generated data

In this example, we randomly generate a graph G that models a set of $n = 20$ individuals and their interactions. We consider two scenarios. In Scenario 1 (see Figure 1a), we assume that there is no information available about the risk level of severe illness from COVID-19 of each individual, nor about the time frames of their social contacts. Hence, we fix the weight of node i, $\delta_i = 1$, for all $i \in \mathrm{V}$. We also assume that all the social contacts have equal time frames and therefore we fix the weight of the edge $\{i, j\}$, $\sigma_{i,j} = 1$, for all $\{i, j\} \in \mathcal{E}$. In Scenario 2 (see Figure 2a), we consider

the same graph G, yet we assume that there is information available about the individual's risk level and time frames of the social contacts. Such information is randomly generated both for the nodes and for the edges. In particular, all the weights are drawn from a uniform distribution between 0 and 1.



(a) Unweighted graph with n = 20 nodes.

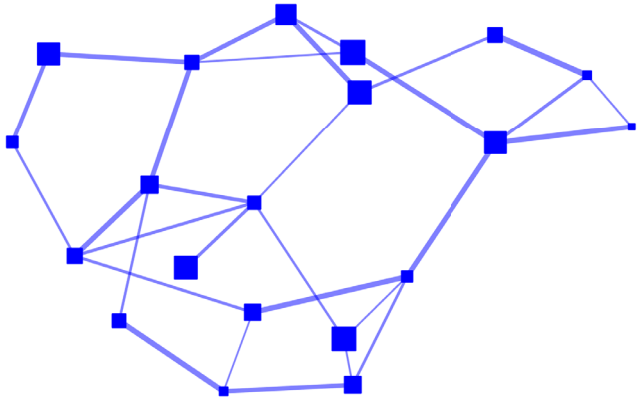


(b) Representation of the 2 clusters created by a single run of Algorithm 1 for the unweighted graph shown in Figure 1a.

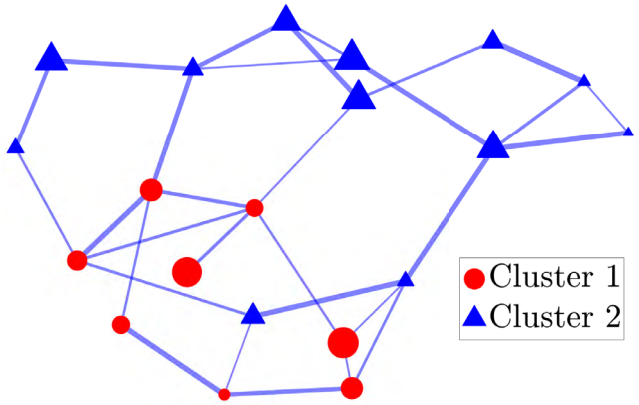
Figure 1. Considered graph and the resulting clustering for Scenario 1

Figures 1b and 2b show the 2 clusters created by a single run of Algorithm 1 for Scenario 1 and Scenario 2, respectively.

Observe that the algorithm does not strictly separate the higher and the lower risk individuals. The clusters made by our algorithm depend on the risk of severe illness but also on the social interaction among individuals.



(a) Doubly-weighted graph with $n = 20$ nodes. The weights for the nodes and the edges are randomly drawn from a uniform distribution between 0 and 1. In the figure, the sizes of the nodes and the widths of the edges are proportional to their corresponding weights.



(b) Representation of the 2 clusters created by a single run of Algorithm 1 for the doubly-weighted graph shown in Figure 2a. In the figure, the sizes of the nodes and the widths of the edges are proportional to their corresponding weights.

Figure 2. Considered graph and the resulting clustering for Scenario 2

4.2. *Example with real data*

In this example, we use data from the CoMix study [42] to generate a doubly-weighted graph G that models a set of $n = 35$ individuals. This study follows households all over Europe, collecting information about their behavioural patterns, measures, and proximity contacts, and how these have varied over time during the course of the COVID-19 pandemic. These results are published for an easier assessment of the spread of the virus, and they maintain

the anonymity of the participants. For this example, CoMix social contact data from Spain were used [43].

From these data, n random participants are selected. CoMix social contact data provides for each participant their number of contacts and the time frame of such contacts. We have further assumed that all the contacts of the selected individuals are within the considered population. We fix the weights of the nodes and the weights of the edges using the information provided by CoMix social contact data as shown in Tables 2 and 3, respectively.

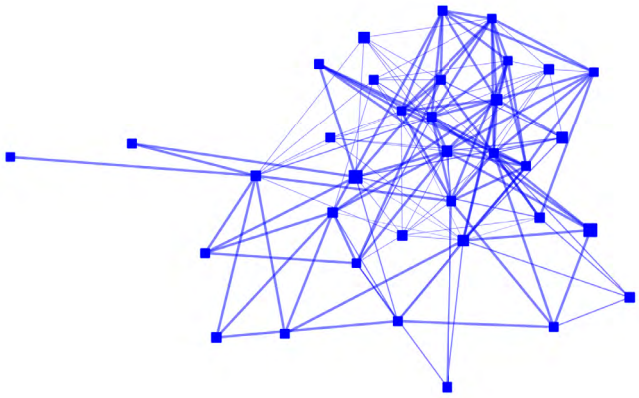
Table 2
Information provided by CoMix about the risk level of severe illness from COVID-19

Age range	Weight of the node
18-29	1/6
30-39	1/5
40-49	1/4
50-59	1/3
60-69	1/2
70-120	1

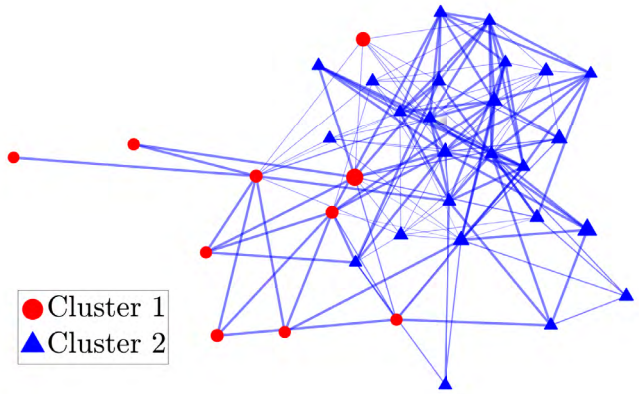
Table 3
Information provided by CoMix about the time frame of the social contacts

Time frame	Weight of the edge
less than 5 minutes	1/16
5-15 minutes	1/8
15-60 minutes	1/4
1-4 hours	1/2
more than 4 hours	1

Figure 3b shows the 2 clusters created by a single run of Algorithm 1 for the considered example.



(a) Doubly-weighted graph with $n = 35$ nodes. The weights for the nodes and the edges are set according to Tables 2 and 3.



(b) Representation of the 2 clusters created by a single run of Algorithm 1 for the doubly-weighted graph shown in Figure 3a.

Figure 3. Considered graph and the resulting clustering for the example with real data

5. Conclusion

In this paper, we have presented a distributed clustering algorithm that groups individuals according to their social contacts and the risk level of severe illness from COVID-19. Once the clusters are made, using a distributed consensus algorithm in each cluster, each individual can know the epidemiological situation of the group they belong to. Such knowledge allows them to take the social distancing measures that correspond to the epidemiological situation of their group. By using this algorithm, the social distancing measures would only affect groups with high risk of infection instead of entire geographical regions, thus reducing the economic

damage. The algorithm is designed so that individuals could know which group they belong to without having knowledge of the rest of the members of the group. Furthermore, there is no central entity with information about the groups because the algorithm only runs at a local level. Groups are created taking into account social contacts and the risk level of severe illness. Since social contacts change continuously and the risk level of severe illness also changes with the vaccination progress, our adaptive algorithm enables the creation of groups according to the information available at the time it is run. Finally, after the computational complexity analysis, we have concluded that our algorithm is sublinear with respect to the population size, which makes it very efficient.

6. References

[1] "Coronavirus disease (covid-19) pandemic," https://www.who.int/emergencies/diseases/novel-coronavirus-2019, accessed: 2021-05-10.

[2] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track covid-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533-534, 2020.

[3] C. Cheng, J. Barceló, A. Hartnett, R. Kubinec, and L. Messerschmidt, "Covid-19 government response event dataset (coronanet v.1.0)," *Nature Human Behaviour*, vol. 4, pp. 756-768, 2020.

[4] N. M. F. *et al.*, "Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand," Imperial College, Tech. Rep., 03 2020.

[5] D. L. Chao, M. E. Halloran, V. J. Obenchain, and I. M. Longini, Jr, "FluTE, a publicly available stochastic influenza epidemic simulation model," *PLOS Computational Biology*, vol. 6, no. 1, 01 2010.

[6] J. K. *et al.*, "Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study," *The Lancet Infectious Diseases*, vol. 20, no. 6, pp. 678-688, 2020.

[7] T. Hale, N. Angrist, E. Cameron-Blake, L. Hallas, B. Kira, S. Majumdar, A. Petherick, T. Phillips, H. Tatlow, and S. Webster, "Variation in government responses to covid-19 (version 7.0)," University of Oxford, Tech. Rep., 09 2020.

[8] J. A. Lewnard and N. C. Lo, "Scientific and ethical basis for social-distancing interventions against covid-19," *The Lancet Infectious Diseases*, vol. 20, no. 6, pp. 631-633, 2020.

[9] N. Fernandes, "Economic effects of coronavirus outbreak (covid-19) on the world economy," *IESE Business School Working Paper*, no. WP-1240-E, 2020.

[10] S. Maharaj and A. Kleczkowski, "Controlling epidemic spread by social distancing: Do it well or not at all," *BMC Public Health*, vol. 12, no. 679, 2012.

[11] "Covid-19 digital rights tracker," https://www.top10vpn.com/research/ investigations/covid- 19-digital-rights-tracker/, accessed: 2021-05-10.

[12] U. Gasser, M. Ienca, J. Scheibner, J. Sleigh, and E. Vayena, "Digital tools against covid-19: taxonomy, ethical challenges, and navigation aid," *The Lancet Digital Health*, vol. 2, no. 8, pp. 425-434, 2020.

[13] M. E. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. van Boven, J. H. H. M. van de Wijgert, and M. J. M. Bonten, "Impact of delays on effectiveness of contact tracing strategies for covid-19: a modelling study," *The Lancet Public Health*, vol. 5, no. 8, pp. 452-459, 2020.

[14] "Tracetogether," https://www.tracetogether.gov. sg/index. html, accessed: 2021-05-10.

[15] H. Cho, D. Ippolito, and Y. W. Yu, "Contact tracing mobile apps for covid-19: Privacy considerations and related trade-offs," arXiv, 2020.

[16] K. Michael and R. Abbas, "Behind covid-19 contact trace apps: The google-apple partner-ship," *IEEE Consumer Electronics Magazine*, vol. 9, no. 5, pp. 71-76, 2020.

[17] S. Jha, L. Kruger, and P. McDaniel, "Privacy preserving clustering," in *Computer Security – ESORICS 2005*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 397-417.

[18] Z. Erkin, T. Veugen, T. Toft, and R. Lagendijk, "Privacy-preserving distributed clustering," *EURASIP Journal on Information Security*, no. 4, 2013.

[19] S. Oliveira and O. Zaiane, "Privacy preserving clustering by data transformation," *Journal of Information and Data Management*, vol. 1, no. 1, 2010.

[20] S. Merugu and J. Ghosh, "Privacy-preserving distributed clustering using generative models," in *Third IEEE International Conference on Data Mining*, 2003, pp. 211-218.

[21] G. Jagannathan and R. N. Wright, "Privacy-preserving distributed k-means clustering over arbitrarily partitioned data," in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*. New York, NY, USA: Association for Computing Machinery, 2005, pp. 593-599.

[22] S. Basagni, "Distributed clustering for ad hoc networks," in *Proceedings Fourth International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN'99)*, 1999, pp. 310-315.

[23] H. Kargupta, W. Huang, K. Sivakumar, and E. Johnson, "Distributed clustering using collective principal component analysis," *Knowledge and Information Systems*, vol. 3, pp. 422-448, 2001.

[24] P. A. Forero, A. Cano, and G. B. Giannakis, "Distributed clustering using wireless sensor networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 707-724, 2011.

[25] E. Januzaj, H. Kriegel, and M. Pfeifle, "Dbdc: Density based distributed clustering," in *Advances in Database Technology - EDBT 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg,2004, pp. 88-105.

[26] M. Jia, Y. Wang, C. Shen, and G. Hug, "Privacy-preserving distributed clustering for electricalload profiling," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1429-1444, 2021.

[27] G. Muniraju, S. Zhang, C. T. M. Banavar, A. Spanias, C. Vargas-Rosales, and R. Villalpando-Hernandez, "Location based distributed spectral clustering for wireless sensor networks," in *2017 Sensor Signal Processing for Defence Conference (SSPD)*, 2017.

[28] A. Bertrand and M. Moonen, "Distributed computation of the Fiedler vector with application to topology inference in ad hoc networks," *Signal Processing*, vol. 93, no. 5, pp. 1106-1117, 2013.

[29] X. Shijie, F. Jiayan, and L. X. Li, "Weighted laplacian method and its theoretical applications," *IOP Conference Series: Materials Science and Engineering*, vol. 768, no. 072032, mar 2020.

[30] W. Chen, Y. Song, H. Bai, C. Lin, and E. Chang, "Parallel spectral clustering in distributed systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 568-586, 2011.

[31] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, pp. 65-78, 2004.

[32] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, ser. NIPS'01. Cambridge, MA, USA: MIT Press, 2001, pp. 849-856.

[33] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, 2000.

[34] D. J. Higham, G. Kalna, and M. Kibble, "Spectral clustering and its use in bioinformatics," *Journal of Computational and Applied Mathematics*, vol. 204, no. 1, pp. 25-37, 2007.

[35] F. Chung and F. Graham, *Spectral Graph Theory*, ser. CBMS Regional Conference Series. Conference Board of the mathematical sciences, 1997.

[36] M. Fiedler, "A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory," *Czechoslovak Mathematical Journal*, vol. 25, no. 4, pp. 619-633, 1975.

[37] B. Hendrickson and R. Leland, "An improved spectral graph partitioning algorithm for mapping parallel computations," *SIAM Journal on Scientific Computing*, vol. 16, no. 2, pp. 452-469, 1995.

[38] G. Dahlquist and A. Bjorck, *Numerical Methods*. Dover, 2003.

[39] J. H. Wilkinson, *The algebraic eigenvalue problem*. Oxford University Press, 1965.

[40] C. Boutsidis, P. Kambadur, and A. A. Gittens, "Spectral clustering via the power method - provably," in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 37. Lille, France: PMLR, 07-09 Jul 2015, pp. 40-48.

[41] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in *Proceedings of the 7th International World Wide Web Conference*, 1998, pp. 161-172.

[42] "The comix study," https://www.uhasselt.be/UH/71795-start/The-CoMix-study, accessed: 2021-09-22.

[43] A. Gimma, K. L. Wong, P. Coletti, and C. I. Jarvis, "Comix social contact data (spain)," Jun. 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5040840

[44] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3, pp. 75-174, 2010.

INPAKTA.EUS

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Mondragon Unibertsitatea

Universidad de Navarra

Université de Pau et des Pays de l'Adour

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Iker

Orkestra
Lehiakortasunerako Euskal Institutua

AGIRRE LEHENDAKARIA CENTER
for Social and Political Studies

Arantzazulab
Gizarte Berrikuntza
Social Innovation

Telesforo Monzon eLab

Deusto
Universidad de Deusto
Deustuko Unibertsitatea

ueu
udako euskal unibertsitatea

BRTA
Basque Research & Technology Alliance

ADItech
Coordinador SINAI
Sistema Navarro de I+D+i

elhuyar
ezagutuz aldatzea

Université Bordeaux Montaigne

Boise State University

Center for Basque Studies

université de Bordeaux

Cenarrusa Foundation for Basque Culture

Soziolinguistika Klusterra
Euskara biziberritzeko ikergunea

Eusko Ikaskuntza
Asmoz ta Jakitez