



Politics of Folk Psychology: Believing what Others Believe[&]

(La política de la psicología folk: creer lo que otros creen)

Uku TOOMING*

Hokkaido University

ABSTRACT: In this paper, I argue that by attributing beliefs the attributer is pushed toward taking a stand on the content of those beliefs and that what stand they take partially depends on the relationship between the attributer and the addressee. In particular, if the addressee enjoys a higher social standing than the attributer, the latter is disposed to adopt the attributed belief, as long as certain other conditions are met. I will call this view the Adoption-by-Attribution model. Because of the non-epistemic influence that derives from the relation of inequality, belief attribution can reinforce the existing unequal power relations and contribute to epistemic injustice.

KEYWORDS: belief; belief attribution, folk psychology; epistemic injustice; testimony.

RESUMEN: *En este artículo sostengo que al atribuir creencias, quien atribuye se ve forzado a tomar una posición respecto al contenido de esas creencias y la posición que toma depende, en parte, de la relación entre quien atribuye y el receptor de la atribución. En particular, si el receptor de la atribución cuenta con una posición social más alta que aquel que atribuye, este último tiene la disposición a adoptar la creencia atribuida, siempre y cuando ciertas otras condiciones se cumplan. Llamaré a esta perspectiva modelo de Adopción-por-Atribución. Debido a la influencia no epistémica que se deriva de la relación de desigualdad, la atribución de creencias puede reforzar relaciones de poder actuales desiguales y contribuir a la injusticia epistémica.*

PALABRAS CLAVE: *creencia; adscripción de creencias; psicología folk; injusticia epistémica; testimonio.*

[&] I am grateful to two anonymous reviewers for insightful and challenging comments on an earlier version of the paper. The research in this paper was supported by JSPS Postdoctoral Fellowship for Research in Japan (Standard), JSPS KAKENHI (19F19762).

* **Correspondence to:** Uku Tooming. Graduate School of Humanities and Human Sciences, Hokkaido University. 4 Chome-1-16 Kita 20 Jonishi, Kita Ward (001-0020 Sapporo, Japan) – uku.tooming@gmail.com – https://orcid.org/0000-0002-4632-7979

How to cite: Tooming, Uku (2021). «Politics of folk psychology: Believing what others believe»; *Theoria. An International Journal for Theory, History and Foundations of Science*, 36(3), 361-374. (https://doi.org/10.1387/theoria.21966).

Received: 2020-08-23; Final version: 2021-03-08.

ISSN 0495-4548 - eISSN 2171-679X / © 2021 UPV/EHU



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

1. Introduction

Attitudes of other people matter to us. It is therefore not surprising that we spend a large chunk of our lives on trying to figure out what other people think and believe. It is a sign of maturity when one is able not to focus too much on others' opinion. However, taking the latter into account is almost inevitable and often reasonable. As we both cooperate and compete with others, our success is dependent on whether we know what they believe.

We wouldn't care about others' beliefs if we didn't have the capacity to attribute beliefs in the first place. Our capacity for belief attribution constitutes one of the main components of mindreading and the latter has been extensively scrutinized in developmental psychology, cognitive science and philosophy. The main debate has revolved around the question of what cognitive mechanisms underlie this capacity. Theory theorists, simulationists and hybrid theorists have each provided their preferred answers to it. Another question has concerned the prevalence of mindreading. While some take for granted that we engage in belief attribution almost whenever we interact with or think about other people, there have been dissenting voices who have argued that in fact belief attribution is often unnecessary for successful intersubjective engagements. Third, there is also a question about the function of mindreading: do we primarily attribute beliefs to explain and predict behavior, to shape the minds of others or to better coordinate with one another? The function could be cashed out in evolutionary terms or in terms of the purposes of agents who attribute beliefs.

In this paper, I do not take a definite stand on these questions. Instead, I will focus on a rather overlooked aspect of belief attribution: its entanglement in relations of social power. In particular, I will propose a hypothesis that when an attributee enjoys a higher social standing than the attributer and some further conditions are met, then the attributer is disposed to adopt the attributed belief. I will call this Adoption-by-Attribution model and present some reasons for why it deserves serious consideration and further empirical investigation.

Although I avoid assuming a particular view about the function of mindreading, the approach I will propose fits well with the mindshaping conception, developed by Tad Zawidzki, and with the regulative view of folk psychology, proposed by Victoria McGeer. Both of them have argued that mental state attributions imply normative expectations to which competent folk psychologists are generally disposed to conform (Zawidzki, 2013; McGeer, 2007). By attributing mental states, we thereby mold the minds of others. We shouldn't think that mindshaping is *the* function of folk psychology¹ but my hypothesis regarding the consequences of belief attribution can be seen as a model of *first-person mindshaping*: at least under circumstances of social inequality, people's belief attributions shape their own beliefs.²

When I say that belief attribution has consequences that depend on the social relationship between the attributer and the attributee, this might bring to mind Vivian Bohl's view, according to which the function of mindreading is to coordinate and shape relationships (Bohl, 2015). The similarities are certainly there, although I do not presuppose Alan

¹ To be fair, neither Zawidzki nor McGeer would say that it's the sole function.

² I'll be focusing on belief attribution but I am fairly confident that, with some tweaks, it can also be extended to desire attributions.

Fiske's relational models theory like she does and I do not suggest that the consequences of belief attribution that I am going to articulate constitute its main function. The focus of this paper is also narrower than Bohl's. First, while she addressed mental state attributions in general, my attention is confined to belief attributions in particular. Second, Bohl's account is also about how social relationships modulate and perhaps even ground our mental state attributions. My account only concerns the downstream effects of already made attributions. Third, if we do presuppose Fiske's theory, then the model that I propose is probably confined to the context of Authority Ranking where the social relationship is represented as hierarchical, and possibly to Market Pricing where it *can be* represented as hierarchical, insofar as different agents are assigned different values (see Fiske, 1992). All in all, it is important to stress that this paper does not pretend to provide a complete view of how social relationships shape the way in which people are affected by their belief attributions. It just suggests a fruitful way of thinking about it.

The paper is structured as follows. In the next section, I am going to argue that attributed beliefs open up a possibility to either agree or disagree with them. This allows me to ask what factors other than epistemic ones affect whether an attributer would agree with the attributed belief and thereby adopt it. In the third section, by relying on studies on testimony, prestige bias and social incentives in belief-formation, I will suggest that whether the attributer is inclined to adopt the attributed belief partially depends on the social rank of the attributee. Finally, in the fourth section, I will consider broader implications of my account regarding power and injustice.

2. *How Others' Beliefs Bear on Us*

The ability to attribute beliefs has various benefits. Among them the most salient one has been the contribution of belief attribution to the explanation and prediction of behavior. Since people's actions are more directly dependent on their beliefs about the world than on the facts, awareness of their beliefs is crucial for understanding and anticipating these actions. The latter are more directly dependent on beliefs because beliefs guide actions even when they do not correspond to facts. It is therefore not surprising that the completion of false belief tasks has been treated as the main indicator of the ability to attribute beliefs (see Wimmer and Perner, 1983; Wellman *et al.*, 2001).

The focus on the context of explanation and prediction guides our theorizing in a particular direction. Since what others do is of immense importance to us and belief attribution helps us get a better grip on it, one might be tempted to think that belief attribution has only instrumental value: if it didn't enable us to make sense of others' actions, we wouldn't engage in it. Even those who think of belief attribution in the context of mindshaping and argue that belief attribution is a practice by which we aim at bringing others' attitudes to conform to our attributions could maintain that mindshaping through belief attribution serves making others' actions predictable and understandable as its ultimate aim.

I am of course not denying that belief attribution has this benefit. However, it seems that others' beliefs matter to us also more immediately, without us having to think about their implications for their possessors' actions. An agent's belief represents the world as being in a certain way and it can be evaluated as either true or false. This evaluative response is open to the attributer independently of their capacity to predict or shape the behavior of

the attributee on the basis of that belief. In fact, it is probably the most immediate response available to the attributer (in Tooming (2017), I flesh this observation out more fully). Others' beliefs are therefore available for evaluation by the attributer. In particular, they are available for agreement or disagreement, and perhaps also for suspension of judgment.³

That these responses are so basic supports the idea that belief attribution has an important role to play in social learning. By having an ability to attribute beliefs to one another, we have potential access to information that others have obtained. If the other person is a reliable epistemic agent, then gaining access to their perspective through belief attribution can have substantial informational benefits. This is because in thinking about others' beliefs, we consider their contents and under favorable circumstances can also infer that those contents are true.⁴ That false beliefs have been the focus in the research on Theory of Mind hides the fact that very often the attributed beliefs are true and besides allowing us to explain and predict their possessors' behavior they also enable us to learn new things about our shared environment. This implies that we then adopt the belief that we have attributed. Whether we open ourselves to this opportunity, however, seems to depend on whether we agree or disagree with the attributee.

Agreement and disagreement have different effects on the relationship between the attributer and the attributee. Disagreement usually produces at least some level of tension between the parties where the other's attitude is seen as problematic; agreement, on the other hand, is conducive to positive regard towards the other. Why do divergent beliefs produce tension? My suggestion is that they constitute a challenge to the attributer. Being aware that another person has a different belief about some subject matter constitutes a challenge to the attributer because the latter needs to ask which of them is right in their belief. One needs to take into account that the other person believes something different, and this can often be quite disturbing. On the other hand, if the attributed belief is compatible with the attributer's own belief, this confirms the attributer's perspective.

Since attributed beliefs can serve either as a challenge or a confirmation, depending on the attributer's evaluation, and thereby open up a possibility of either agreeing or disagreeing with them, from the attributer's standpoint they are functionally similar to assertions. When an agent asserts something, they take the world to be in a certain way and make it possible for their audience to either agree or disagree with them. Both the attributed belief and received assertion have treated-as-true-contents which the attributers and receivers can agree or disagree with. Of course, there are obvious differences between responding to an assertion and responding to an attributed belief. In asserting, one explicitly and consciously presents a claim to the audience and makes it available for agreement or disagreement. In the case of attributed belief, on the other hand, the attributer's responses are independent of the agency and awareness of the attributee.

But despite these differences, it is still noticeable that the basic response profile of agreement or disagreement is shared between assertions and attributed beliefs. Both assertions that one disagrees with and divergent beliefs function as challenges to the perspective of the audience or the attributer, respectively, and both assertions that one agrees with and

³ On the suspension of judgment as a distinct attitude, see Friedman (2013).

⁴ That thinking about others' beliefs involves thinking of what these beliefs are about has also been stressed by Jane Heal in her co-cognition account (Heal, 1998).

compatible beliefs function as a confirmation. These parallels between attributed beliefs and assertions suggest that understanding the consequences of disagreement and agreement with the latter could help us understand the consequences of those evaluative responses to the former.

In fact, responding to assertions is arguably derivative with respect to responding to beliefs, in the sense that in agreeing or disagreeing with an assertion we are agreeing or disagreeing with the belief that it expresses. Arguably, since comprehending an assertion involves attributing a belief that it expresses, *every case* of comprehending an assertion is already a case of attributing a belief and *every case* of agreeing with that assertion already involves agreeing with the corresponding belief (and adopting it if the hearer did not already have it). Therefore, instead of talking about the similarities between responding to assertions and responding to attributed beliefs, we can as well just talk about the similarities between responding to beliefs that are attributed on the basis of explicit assertion and beliefs that are attributed by other means.

Given that there is plenty of literature on the psychology of forming beliefs on the basis of testimony and testimony comes in the form of assertions, I suggest that this literature provides a source which can be tapped into in order to better understand the conditions under which belief-attributions contribute to the belief-formation of the attributers. Therefore, in the next section, I will explore how the data on how people process testimony can illuminate our understanding of how people process attributed beliefs.

It could be insisted that we should not rely on the data on how people respond to belief attributions on the basis of assertions to draw conclusions about factors that affect responses to belief attributions across the board because the belief attributions on the basis of assertions are relevantly different from the rest. It is correct to point out, for instance, that in asserting, one can try to intentionally deceive the hearer about one's beliefs, while when a belief attribution is made on the basis of nonverbal behavior, for instance, the attributee is not in a position to deceive in the same way.⁵

I am a bit doubtful if the difference between the two kinds of cases is as radical as it may seem, though. After all, one can also deceive by acting non-verbally in a way that is suggestive of a belief that one actually does not have. But even if we grant that deception is easier in the case of belief attribution on the basis of explicit communication, this does not give a reason to think that the factors that affect whether the attributed belief is agreed or disagreed with are substantially different in two kinds of cases. Easiness to deceive influences which beliefs are attributed or not, it does not affect whether already attributed beliefs are agreed or disagreed with. Whether the attribution is true or not, justified or not, is not relevant in the present context.

What would make a relevant difference would be a discovery that influences on assertion-based belief attributions do not bear on belief attributions that are not based on assertions. I see no reason for claiming, however, that somehow the latter are immune to those factors. Whether we agree with another person's belief depends on whether we find that person credible, and credibility assessment seems to figure in our responses to attributed beliefs both in the case of verbal and non-verbal behavior. It is thus a good bet to think that the data on the former also bears on the latter.

⁵ I thank an anonymous reviewer for stressing that difference.

Before going on, it is important to stress that what will be said in what follows about the factors in responding to attributed beliefs, is provisional and amenable to further empirical investigation. The aim is to propose a hypothesis regarding a non-epistemic factor (considerations regarding the social relationship between attributers and attributees) that plausibly plays a role in biasing the way in which attributers respond to the beliefs that they have attributed. Aside from the data on testimony, I will also appeal to some other research.

3. *Adoption-by-Attribution*

It is a truism that we form many of our beliefs on the basis of others' assertions. In fact, it can also be argued that others' assertions give one a distinctive reason to believe their contents, a reason that isn't reducible to any evidential consideration regarding the content of the assertion (cf. Moran, 2018). In the context of this paper, I don't need to assume this, however. What is important is that, even if they are actually mistaken in assuming this, people do *treat* others' assertions as providing a reason to believe. There is a general tendency, formed already in early childhood, to believe what other people say. This is indicated by the fact that 12-13 month old infants learn names of novel objects from agents whose prior reliability they have had no opportunity to evaluate (Bloom, 2002; Waxman and Braun, 2005). Also, a testimony about an object constrains children's attention to the referred features, indicating that they take others' assertions to provide information that closes off further exploration of the object (Bonawitz *et al.*, 2011). Infants' trust extends even to deceptive testimony (cf. Jaswal *et al.*, 2010). Or take studies by Vanderbilt and colleagues who showed that 3- and 4-year-olds accepted the testimony from a previously inaccurate agent, at least when there was no conflicting testimony available (Vanderbilt *et al.*, 2014). Even speakers' earlier admission of ignorance doesn't hinder children from accepting it (Kushnir and Koenig, 2017). It seems, then, that people take testimony to provide a reason to believe what is said and thereby adopt the belief that they take the testimony to convey.

It is another truism, however, that people are not so gullible that they adopt any belief that they have attributed to others on the basis of testimony. Thus, even if the fact that a person asserts that *p* is taken by someone to be a reason to believe that *p*, they need not, and often do not, adopt that belief themselves. There are mechanisms of epistemic vigilance which enable people to avoid taking on beliefs of others when this leads to sub-optimal results or when this is epistemically harmful (Sperber *et al.*, 2010). These mechanisms take shape already in early childhood. Whether one is motivated to adopt the attributed belief hinges upon the seeming credibility of the one who believes. But since credibility cannot be immediately perceived, one has to rely on cues which indicate it. What are those cues?

The literature on how children process verbal testimony provides useful information in that regard as well. First, from early on in the development, credibility of a claim (and thereby a belief that the claim is taken to express) is evaluated in terms of epistemic considerations. For instance, children trust verified claims more than unverified claims (Butler *et al.*, 2018). Credibility of a piece of testimony is also a function of the perceived credibility of the testifier: children's trust in informants is selective, tracking the past reliability with the particular domain they are informed about (Harris and Corriveau, 2011). In gen-

eral, competence in a domain contributes to the credibility regarding that domain (Harris *et al.*, 2018). We can assume that our evaluation of an attributed belief thus depends to a large extent on epistemic considerations, and we generally tend to adopt that belief when we deem it credible on the basis of those considerations. When one has a reliable informant who can be trusted, it is possible to trust the beliefs that have been attributed to her.

However, not only epistemic considerations guide the way in which we evaluate others' credibility. Although the latter should be a function of epistemic competence, features of the informant's character that are not directly epistemically relevant also affect whether we believe what we take the other to believe. Plenty of research suggests that other people's social standing and personality affect whether they are taken to be credible and trustworthy and whether their testimony is accepted or not (for an overview, see Harris *et al.*, 2018). For instance, already children are more inclined to accept testimony from ingroup members (Levy, 2019), from agents who they are familiar with (Corriveau and Harris, 2009) and who seem more confident (Tenney *et al.*, 2011). In addition, they show preference toward the consensus opinion (Corriveau *et al.*, 2013).

People also show doxastic preference toward agents who others in their group have preferred and continue to prefer as informants (Chudek *et al.*, 2012). This is what has been called "prestige bias". Joseph Henrich has stressed the role of prestige in cultural transmission, among other factors (Henrich 2009, p. 245). People tend to prefer highly prestigious individuals when looking for models to learn from (Henrich and Gil-White, 2001). Importantly, prestige bias is not limited to how testimony is processed, it is a bias that characterizes cultural learning more broadly. Learning from prestigious individuals is an effective strategy because prestige generally predicts epistemic competence. Evaluating epistemic competence in itself is costly, so relying on prestige provides a faster way to track it (Atkisson *et al.*, 2012, p. 521). It is thus natural to think that it extends to attributed beliefs as well, so that the beliefs of prestigious individuals have an advantage of spreading in a population. Although this may quite often result in adopting false beliefs, it is in general a useful strategy.

How prestige affects whose opinions in the group are adopted was investigated in a recent study by Brand *et al.* (2020). Participants in their experiment were assigned to three experimental conditions – Control, Prestige, or Success. They were asked quiz questions in two rounds. In Round 1, they were given one cue to use if they decided to copy someone else's (demonstrator's) opinion, in Round 2, they could choose between two cues. In the Prestige condition, the information in Round 1 concerned the scores of the demonstrators up to that point, indicating their differential success. In Round 2, the choice was between the Player ID of the demonstrators (irrelevant information) or information regarding how often different demonstrators were copied, i.e., the difference in prestige. In the Control condition, on the other hand, although the choice in Round 2 was the same, the information in Round 1 concerned Player IDs, which made information about the frequency of copying unreliable in Round 2. In the Success condition, information given in Round 1 was the same as in the Prestige condition, but in Round 2, the choice was between information regarding the success of demonstrators up to that point and the difference in prestige.

As it turned out, in Round 2, the participants preferentially copied the most prestigious demonstrator's opinion both in the Prestige and Success condition, although in the Success condition, direct success cues were preferred over prestige cues. In contrast, in the Control condition, they did not copy the most prestigious demonstrator in Round 2. Since

in the latter, the information gained in Round 1 was not reliable, so was information about the frequency of copying unreliable in Round 2. These data indicate that people are disposed to rely on prestige in adopting the ideas of others but only when the prestige is associated with past reliability and when direct information about epistemic reliability is not available. Perceived prestige has an influence on whether one is disposed to adopt the attributed belief, but this influence is not immune to interfering factors like evidence of unreliability.

The social relationship between the attributer and the attributee therefore matters for whether the attributer adopts the attributed belief, although in a quite nuanced way. Relatedly, there are reasons to think that also perceived dominance of the attributee over the attributer makes the attributed belief more attractive to the latter. Indirect evidence for this comes from a study which suggests that 3-5 year olds prefer the testimony of a dominant over a subordinate, where dominance was measured either by the physical strength or the ability to impose goals on the other (Bernard *et al.*, 2016, p. 315). I take it that explicit endorsement of testimony is indicative of forming a belief that the content of testimony is true.⁶ Of course, prestige and dominance are not the same: one is a type of social rank that is partially sustained by competence, while the other need not be based on competence but is instead upheld by a threat of force. However, both constitute ways in which an agent can enjoy a higher social status in the community, and both predict social influence, measured by the degree to which others align their thoughts and opinions with those of the agent (Cheng *et al.*, 2013, p. 111).

Let's call the model of adopting attributed beliefs because of the higher perceived social status of the attributee Adoption-by-Attribution model. On an opposing, purely rationalist, picture, the adoption of a belief is motivated by considerations pertaining to whether the proposition in question is supported by evidence and how well it coheres with the ascriber's overall system of belief. The view that I support here diverges from such a rationalist view by acknowledging the possibility of non-evidential "brute" belief-formation. The adoption of a belief may be mediated by nothing more than the authority of the other.

Aside from the data on testimony and on prestige bias there is a further reason why Adoption-by-Attribution is plausible. Daniel Williams (2021) has recently argued for the view that human belief formation is responsive to social incentives and not just epistemic considerations. According to Williams, this reward-sensitivity explains why people are prone to confabulation, positive illusions and identity protective cognition, all of which push people toward showing themselves in the positive light to others. If Williams is right about the role of social incentives in belief formation, then this also increases the plausibility of Adoption-by-Attribution. Assuming that our belief formation is sensitive to social rewards and pleasing one's superior is a prime example of such a reward, it is expected that people are disposed to adopt the beliefs of their superiors. Williams's hypothesis offers a plausible explanation as to why Adoption-by-Attribution could have developed: it developed because forming beliefs through it was socially rewarded.

⁶ It should be stressed, though, that a mere fact of dominance does not supersede weak arguments in making a person credible (cf. Castelain *et al.*, 2016). This mirrors the way in which prestige influences belief-formation: direct evidence of reliability or unreliability is prioritized over prestige.

Regarding Adoption-by-Attribution, some qualifications are in order. First, assuming that Adoption-by-Attribution is motivated by social incentives, it might not apply to those attributed beliefs which would not or could not be expressed to others, were one to adopt them. After all, in the case of such beliefs, there is no social incentive for adopting them, given that others would not be able to know that one has adopted them.

Second, it is important to stress that Adoption-by-Attribution doesn't imply that people start believing whatever they think that others with higher social status believe. It is a dispositional notion according to which people are motivated to adopt the beliefs of their social superiors, without entailing that they necessarily do so. As acknowledged above, people are generally not motivated to adopt beliefs that they take to be completely unreasonable. It should also be stressed that a mere superior position of the attributer is not sufficient for the attributer to adopt the attributed to belief. As the study by Brand *et al.* indicated, the disposition to adopt the beliefs of one's superior can be undercut by information about the superior's unreliability about the domain in question. However, as long as the facts about unreliability are unavailable to the attributer, then they presumably continue to rely on Adoption-by-Attribution. Also, given the unequal relationship between the attributer and attributee, the attributee can also contribute to that unavailability by using their social capital to make salient those occasions in which their beliefs have been correct while turning the attributer's attention away from cases in which they have been wrong.

Third, there is a difference between a situation wherein the attributer has an existing belief which is inconsistent with the belief that is attributed and a situation wherein the attributer does not yet have any opinion on the matter.⁷ The belief attribution affects the attributer in different ways in those two cases. Plausibly, in the first case, the attributer is less prone to adopt the attributed belief. For instance, when one has acquired a belief on the basis of perception under normal circumstances, then the awareness of another person's divergent belief does not push one to adopt the latter (Clément 2010). It is thus not the case that Adoption-by-Attribution indiscriminately applies to all situations wherein a belief is attributed to a social superior. That being said, Adoption-by-Attribution still provides a quick and easy path to new beliefs when the relevant conditions are met.

Fourth, we should also keep in mind that Adoption-by-Attribution need not apply to all people who are perceived as socially prestigious or dominant. As noted above, literature on testimony indicates that people show preference to informants who belong to their ingroup (Chen *et al.*, 2013). Applying this observation to attributed beliefs, together with Adoption-by-Attribution, this seems to suggest that one is inclined to adopt prestigious individuals' beliefs only when they belong to one's ingroup. Adoption-by-Attribution therefore does not apply to attributed beliefs of prestigious agents across the board.

Before moving on, I should also briefly consider an account of belief-formation which might seem to make the Adoption-by-Attribution model theoretically much less interesting. A more radical view than Adoption-by-Attribution would claim that if an agent comprehends a proposition, then they automatically believe it. Only after this automatic response it is possible to doubt it. This "Spinozist" model was suggested by Daniel Gilbert

⁷ I disregard situations wherein the beliefs of the attributer and the attributee already match because one can't presumably speak about adopting a new belief in such a case.

(1991) and has recently been defended by Eric Mandelbaum who has brought plenty of empirical evidence in favor of the idea that a proposition that is represented is automatically believed (Mandelbaum, 2014). Spinozism may seem to be in tension with Adoption-by-Attribution because the latter assumes that one starts to believe a proposition only when certain conditions are met, namely, that it is believed by the attributee who enjoys higher social status. Mere comprehension of a proposition does not suffice.⁸

I think, however, that Adoption-by-Attribution merits theoretical interest even if we accept the Spinozist hypothesis. We can respond by claiming that Adoption-by-Attribution does not characterize automatically held beliefs but attitudes which express subjects' conscious assent to a proposition. Even Spinozists postulate more reflectively held cognitive attitudes which are distinct from beliefs that are formed automatically by just tokening a representation. For instance, Jake Quilty-Dunn distinguishes between automatic and endorsed beliefs: while the former are described by the Spinozist model, the latter are what subjects take themselves to have and are not the result of mere comprehension of a proposition (Quilty-Dunn, 2015, p. 279). Adoption-by-Attribution is primarily about the latter, not the former. It characterizes beliefs endorsed by the subject in which case the formation of such an endorsement is not automatic but sensitive to the social standing of the other, among other factors.

4. *Belief Attribution and Social Power*

I will conclude this paper by elaborating on the concept of social power and its relation to Adoption-by-Attribution. As a preliminary definition, let's think of social power as a capacity of agents to make a difference to the way in which other agents act and think, independently of whether they agree to this. We can also add two qualifications. First, social power is not just a power to do something (for such a conception, see Hobbes, 1996, p. 58), but a power over someone (Dahl, 1957, p. 202). Second, social power is dispositional, which means that it is manifest only when certain conditions are met (Lukes, 2005, p. 69). Accordingly, prestigious and dominant individuals have social power because of having those personal qualities which constitute the capacity to influence other people. And, since power is dispositional, one need not exercise this capacity in order to have it.

We can now ask: to what extent is the social power of the attributee modulated by others' belief attributions, in the sense of enhancing their capacity to influence attributers? I think that there is an important possibility here. If the attributer is inclined to adopt the latter's belief only because of their unequal relationship, the social power of the attributee is then consolidated by the fact that one is disposed to adopt another's belief. The attributee's influence on the attributer extends to the latter's attitudes and, because of this, the attributee gains in social dominance. Belief attributions that activate Adoption-by-Attribution can perpetuate already existing power relations.⁹

⁸ It is an open question if Spinozism is the correct view of believing. Arguably, people do not so easily believe a content when it is highly relevant for them (cf. Sperber *et al.*, 2010).

⁹ In fact, it can be argued that the ability to control others' attitudes is the highest form of social power (Lukes, 2005, p. 27).

From this description, since the capacity to influence others tends to be associated with intentional agency, one might get an impression that the process of perpetuating social power is under the intentional control of the attributee. Given that Adoption-through-Attribution is a process that neither the attributer nor the attributee intentionally control, this may seem rather implausible. However, although I stated in the beginning of this section that social power consists in a capacity to make a difference to what others think and do, this capacity itself need not be underwritten by intentional control. There are other ways to conceptualize power than in terms of individual agency. According to the systemic conception, social power, including the power of individual agents, is grounded in ways in which a social organization structures one's capacity to act and think (Haugaard, 2010, p. 425).¹⁰ We do not need to think of social organization in a narrowly institutional way. Among other things, a social organization concerns the distribution of prestige and dominance in a community which underwrites one's capacity to control others, both intentionally and unintentionally. According to the present picture, this distribution, in part through Adoption-by-Attribution, prioritizes some individuals when it comes to believing what others are taken to believe. Adoption-by-Attribution therefore contributes to the way in which some people's capacity to (unintentionally) influence others' thoughts is shaped by the social organization.

An appeal to the systemic conception of power might leave an impression that the intentional agency of the attributees has no role to play in affecting the power dynamic in question. This is not entirely true, however, because there is still some room for *indirect* intentional control. For instance, what the attributee does sometimes control to an extent is whether the attributer can keep track of the attributee's actual epistemic credentials. As pointed out in the previous section, prestige usually originates in some actual competence, but it is often not easy to determine whether it still reliably signals such competence in the present.¹¹ The prestigious attributee can mislead the attributer into thinking that the attributee still possesses the competence, even if they do not. And when the attributee manages to do this, they manage to maintain their social power.

Finally, what is the moral status of this kind of power dynamic that Adoption-by-Attribution facilitates? One can say that the inclination to adopt the beliefs of the more powerful is conducive to a form of epistemic injustice in which case a member of a subordinate group is hindered from fully exercising their epistemic powers due the excessive credibility that is attributed to the members of a dominant group. This kind of over-estimation of the other's credibility (and under-estimation of their own) puts the attributer in an epistemically inferior position without proper justification. Compare this to Miranda Fricker's conception of testimonial injustice which involves giving speakers less credibility due to

¹⁰ I am not assuming here that the systemic conception provides an overarching theory of social power, but only that it illuminates one aspect of it, namely, it lets us see how power is not necessarily upheld by intentional agency.

¹¹ This is important to keep in mind also because the reason why Adoption-by-Attribution developed is presumably that it benefited the attributers: it provided them with a quick way to acquire reliable information and useful skills from more competent agents (Henrich and Gil-White, 2001). However, in circumstances where the link between prestige and competence is broken but the attributers do not know that, it primarily benefits the attributee, although even then it is plausible to think that the attributer can reap some social benefit by aligning their beliefs with the highly ranked attributee.

prejudice (Fricker, 2007). Transferring her idea from the context of testimony to that of belief-attribution, the kind of epistemic injustice that can be inflicted by Adoption-by-Attribution involves giving excessive credibility to the attributee due to their social rank. Fricker herself, at least in her 2007 book, doesn't take attribution of excessive credibility to be a form of injustice because it doesn't involve not giving sufficient respect to the speaker as an epistemic subject and also because credibility is not a scarce good in which case one person's abundance would imply another person's lack (Fricker, 2007, p. 20). However, a number of authors has argued that also excessive credibility can constitute injustice (see Medina, 2010; Davis, 2016). In the present case, the attributed excessive credibility is clearly epistemically harmful to the attributer because this affects their belief-formation in a way that is damaging to their autonomy and epistemic agency.

What could be done to alleviate this kind of injustice? Again, although Adoption-by-Attribution is not itself a directly controllable process, there are various ways to undermine its influence indirectly. From the perspective of less powerful attributers, one possible strategy is to try to give less credence to propositions that are believed by more powerful attributees. This might reduce the effect of Adoption-by-Attribution on their belief-formation.¹² From the perspective of more powerful attributers, if they want to avoid having such an effect on others, they should be mindful of the effect of their authority on the belief-formation of less powerful interpreters. They should acknowledge that agents of lower social status are doxastically vulnerable to adopting their beliefs in virtue of their perceived status and they should try to interfere when another person has adopted their belief through Adoption-by-Attribution.

However, since one cannot always check what beliefs others have attributed and/or adopted and any intervention with respect to others' beliefs might still activate in them Adoption-by-Attribution, a more fruitful strategy is to try to change the social conditions in a way that would make its activation less likely. One way to contribute to this is to make the society more egalitarian because Adoption-by-Attribution is activated only under circumstances of inequality. This is where doxastic vulnerability that derives from belief-attribution can become a political challenge.

REFERENCES

- Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology*, 10(3), 519-534.
- Bernard, S., Castelain, T., Mercier, H., Kaufmann, L., Van der Henst, J. B., & Clément, F. (2016). The boss is always right: Preschoolers endorse the testimony of a dominant over that of a subordinate. *Journal of Experimental Child Psychology*, 152, 307-317.
- Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: MIT press.
- Bohl, V. (2015). We read minds to shape relationships. *Philosophical Psychology*, 28(5), 674-694.

¹² Compare this to Shannon Spaulding's suggestion regarding the way in which a person from a dominant group should react to peer disagreement with a subordinate. According to Spaulding, since we have self-serving biases and tend to give less credibility to the assertions of members of subordinate groups, we should treat the latter as at least moderately epistemically superior to us, in order to counteract those biases (Spaulding, 2018, p. 87).

- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120(3), 322-330.
- Brand, C. O., Heap, S., Morgan, T. J. H., & Mesoudi, A. (2020). The emergence and adaptive use of prestige in an online social learning task. *Scientific Reports*, 10(1), 1-11.
- Butler, L. P., Schmidt, M. F., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology*, 176, 73-83.
- Castelain, T., Bernard, S., Van der Henst, J. B., & Mercier, H. (2016). The influence of power and reason on young Maya children's endorsement of testimony. *Developmental Science*, 19(6), 957-966.
- Chen, E. E., Corriveau, K. H., & Harris, P. L. (2013). Children trust a consensus composed of outgroup members—but do not retain that trust. *Child Development*, 84(1), 269-282.
- Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, 104(1), 103-125.
- Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, 33(1), 46-56.
- Corriveau K. H., & Harris P. L. (2009). Choosing your informant: weighing familiarity and recent accuracy. *Developmental Science*, 12, 426-437.
- Corriveau, K. H., Kim, E., Song, G., & Harris, P. L. (2013). Young children's deference to a consensus varies by culture and judgment setting. *Journal of Cognition and Culture*, 13(3-4), 367-381.
- Dahl, R. (1957). The concept of power. *Behavioral Science*, 2, 201-215.
- Davis, E. (2016). Typecasts, tokens, and spokespersons: A case for credibility excess as testimonial injustice. *Hypatia*, 31(3), 485-501.
- Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, 99(4), 689-723.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Friedman, J. (2013). Suspended judgment. *Philosophical Studies*, 162(2), 165-181.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107-119.
- Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, 69, 251-273.
- Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1179-1187.
- Haugaard, M. (2010). Power: A 'family resemblance' concept. *European Journal of Cultural Studies*, 13(4), 419-438.
- Heal, J. (1998). Co-cognition and off-line simulation: Two ways of understanding the simulation approach. *Mind & Language*, 13, 477-498.
- Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, 30(4), 244-260.
- Henrich, J. & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22(3), 165-196.
- Hobbes, T. (1996 [1641]). *Leviathan*. Oxford World's Classics. Oxford University Press.
- Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, 21(10), 1541-1547.
- Kushnir, T., & Koenig, M. A. (2017). What I don't know won't hurt you: The relation between professed ignorance and later knowledge claims. *Developmental psychology*, 53(5), 826-835.
- Levy, N. (2019). Due deference to denialism: explaining ordinary people's rejection of established scientific findings. *Synthese*, 196, 313-327.
- Lukes, S. (2005). *Power: A radical view*, 2nd expanded edition. London: Macmillan.
- Mandelbaum, E. (2014). Thinking is believing. *Inquiry*, 57(1), 55-96.

- McGeer, V. (2007). The regulative dimension of folk psychology. In D. Hutto & M. Ratcliffe (Eds.), *Folk psychology re-assessed* (pp. 137-156). Springer Netherlands.
- Medina, J. (2011). The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary. *Social Epistemology*, 25(1), 15-35.
- Moran, R. (2018). *The Exchange of Words: Speech, testimony, and intersubjectivity*. Oxford: Oxford University Press.
- Quilty-Dunn, J. (2015). Believing in perceiving: Known illusions and the classical dual-component theory. *Pacific Philosophical Quarterly*, 96(4), 550-575.
- Spaulding, S. (2018). *How we understand others: Philosophy and social cognition*. Routledge.
- Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, 25(4), 359-393.
- Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology*, 47(4), 1065-1077.
- Tooming, U. (2017). Beliefs and desires: From attribution to evaluation. *Philosophia*, 45(1), 359-369.
- Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science*, 17(3), 443-451.
- Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, 95(3), B59-B68.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of Theory-of-Mind development: the truth about false belief. *Child Development*, 72, 655-684.
- Williams, D. (2021). Socially adaptive belief. *Mind & Language*, 36(3), 333-354.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.
- Zawadzki, T. W. (2013). *Mindshaping: A new framework for understanding human social cognition*. Cambridge, MA: MIT Press.

UKU TOOMING is a JSPS International Research Fellow at the Graduate School of Humanities and Human Sciences at Hokkaido University. His main research interests are in philosophy of mind, aesthetics and epistemology.

ADDRESS: Graduate School of Humanities and Human Sciences, Hokkaido University. 4 Chome-1-16 Kita 20 Jonishi, Kita Ward, 001-0020, Sapporo, Japan. Email: uku.tooming@gmail.com
ORCID: <https://orcid.org/0000-0002-4632-7979>