

Effects of Scientific Information Format on the Comprehension Self-Monitoring Processes: Question Generation

Vicente Sanjosé*, Tarcilo Torres**, and Carlos Soto**

*University of Valencia (Spain), **University of Antioquia (Medellín, Colombia)

Abstract

Generating questions is a regulatory action associated with self-monitoring processes in comprehension tasks: subjects can ask information seeking questions to solve comprehension obstacles. A sequence of two related experiments were conducted to trigger, classify and analyse questions asked under different conditions: reading a text about experimental scientific devices operating, watching these devices in a DVD and manipulating them in the LAB. Students' information seeking questions were classified using a simple taxonomy. Taking into account the multimedia learning principles, the advantages of realistic animations for understanding time-depending processes and the effect of the procedural-motor activity, were expected students would ask different questions under each of the aforementioned conditions. Results confirmed the expectations: the reading condition triggered more questions addressed to describe the entities while the watching and manipulation conditions stimulated more causal questions. In addition, the effect of prior knowledge on questions including scientific concepts was analyzed.

Keywords: Science education, self-regulation, question generation, experimental devices.

Resumen

Formular preguntas es una de las acciones asociadas con procesos metacognitivos de regulación durante tareas de comprensión: los sujetos pueden generar preguntas destinadas a obtener información para salvar obstáculos de comprensión. Se realizaron dos experimentos para estimular, clasificar y analizar las preguntas de los estudiantes en tres condiciones diferentes: leer sobre el funcionamiento de dispositivos científicos experimentales, observarlos en un DVD o manipularlos en el Laboratorio. Las preguntas fueron clasificadas usando una taxonomía sencilla. Teniendo en cuenta los principios multimedia de aprendizaje, las ventajas de imágenes realistas animadas para el aprendizaje de procesos dependientes del tiempo y el efecto de la actividad procesual-motora, se predijo diferente distribución de preguntas en las tres condiciones consideradas. Los resultados confirmaron lo esperado: la lectura de textos estimuló más las preguntas destinadas a describir las entidades, mientras las condiciones de observar y manipular estimularon más las preguntas causales. Además, se analizó el efecto del conocimiento previo sobre las preguntas que incluyen conceptos científicos.

Palabras clave: Didáctica de las ciencias experimentales, autorregulación, generación de preguntas, dispositivos experimentales.

Acknowledgement: Research financed by the University of Valencia (Professors Mobility Program) and the University of Antioquia (Sustainability Program 2011-2012).

Correspondence: Vicente Sanjosé, Science Education, University of Valencia. Avda Tarongers 4, 46022-Valencia. E-mail: vicente.sanjose@uv.es

Introduction

Teachers agree about the importance of student questions for deep learning (Chin, 2002; Chin & Osborne, 2008; Dillon, 1988; Flammer, 1981; Watts, Gould, & Alsop, 1997). In general, regulatory activities are good predictors for academic success (García-Ros & Pérez-González, 2011; Wang, Haertel, & Walberg, 1993) and asking a question is a self-regulation activity.

In academic contexts the most important kind of questions are those directed at solving a lack of knowledge, or *Information Seeking Questions* (ISQ). Our aim is to study how different information formats influence ISQ generation. Specifically, we want to know whether question generation varies when students read about, visualize or handle experimental scientific devices. According to multimedia learning principles (Mayer, 2001, 2005), the format in which information is presented would affect the mental representation students' attempted to build. Thus, the obstacle-goal model about question generation (Otero, 2009) predicts that different types of questions would be asked when the operation of scientific devices was presented to the subjects in different formats: reading about the operation of scientific devices (Text condition onwards), watching on a video the same devices while they are operating (DVD condition), or han-

dling these devices in the laboratory (LAB condition).

Questions and the information format

Compared to information presented only in text format, integrating visual information of images, graphics and pictures with text can improve comprehension (Schnotz, 2005). Moreover, according to the Cognitive Load Theory (Chandler & Sweller, 1991), seeing a process or a procedure may result in a reduction of the cognitive load compared to a situation in which the process has to be reconstructed from a series of still (static, as opposed to dynamic) images or by reading a text; hence, animations (dynamic images) could one help to understand scientific phenomena more than static images could. Höffler and Leutner (2007) found a medium-sized overall advantage of instructional animations over static pictures and more substantial effect sizes when the animation was highly realistic.

In addition, handling real-life devices could produce differences in learning outcomes with respect to reading about or watching them. Some recent approaches to cognition claim that the conceptual system is grounded in the brain's modality-specific systems, in the body and in the environment (Barsalou, 2012). Conceptual representations are modal, i.e. representations of concepts strongly depend

on the perception modality (visual, handling, etc). Thus, some differences in comprehension should appear when students can handle real world devices, only visualize them, or they read about them. Ferguson and Hegarty (1995) found some advantages from handling real devices with respect to visualizing still pictures or diagrams, in problem-solving transfer tasks.

Questions and previous knowledge

The influence of previous knowledge on the questions generated has been investigated (Otero & Grasser, 2001; van der Meij, 1990) and two hypotheses have been considered: (a) lack of knowledge; (b) cognitive conflict. According to the former, a question is generated when the subject does not know the particular information he/she needs to know. Nevertheless Miyake and Norman (1979) showed that lack of knowledge cannot explain the diversity of generated questions. Graesser and Olde (2003) studied questions formulated about a broken device by participants with different levels of expertise, and they found that the deeper the comprehension the higher the number of adequate questions for repairing the device. These results corroborate the cognitive conflict hypothesis claim: questions depend on inconsistencies or incompatibilities between external

information and the subject's previous knowledge.

Moreover, the nature of the questions asked will depend on the particular knowledge the subject has got. When a student has low scientific knowledge about, for instance, Newton's three laws of the dynamics, it is difficult for him/her to ask a question such as "What is the resultant force causing this acceleration towards the upper side of the slope?" Instead, he/she will probably ask "Why does the object move towards the upper side of the slope? The second formulation refers to common knowledge and it is associated with the construction of the mental representation called *situation model*, whereas the first one is associated with the *scientific model*.

Questions and inferences in science comprehension

According to the *obstacle-goal* model (Otero et al., 2008) a question is generated when a subject finds a comprehension obstacle in his/her way to the desired goal, i.e. the elaboration of certain mental representation able to integrate the provided information. Science comprehension implies the construction of the high level mental representations (Gangoso, 2004; Greeno, 1989) and the activation of previous knowledge and inference generation are important cognitive processes in the construc-

tion of these mental representations. Making inferences may be the most important cognitive activity for science comprehension (Graesser & Zwaan, 1995). According to the Otero's model, ISQ generated in science comprehension tasks should be associated to specific obstacles found by students: the intended but failed inferences (Otero, Ishiwa, & Sanjosé, 2008) they try to build.

Trabasso and Magliano (1996) identified three broad categories of inferences generated in the process of the conscious understanding of a narrative text: associative, explanatory and predictive. *Associations* provide information about the features, properties, relationships and descriptive details of the entities (actors, objects, actions and events). *Explanations* provide reasons about why something occurs. Lastly, *Predictions* are forward-oriented encompassing the consequences of actions or occurrences, and allow for one to anticipate events.

This simple taxonomy of inferences has been extended to the field of expository, scientific information, and the corresponding taxonomy of questions has been proposed (Otero et al., 2008). First, the Q1 questions are linked to failed *associations* and frequently formulated by using *who*, *what*, *how*, *when* and *where*, addressed to the need of adequately representing the entities of the system under consideration, as well as their

properties. The *explanation* questions, Q2, or *why* questions, are addressed to justify or explain why entities are the way they are. *Predictive* questions (Q3) are generated from *if-then* failed inferences, and they are aimed to foresee consequences as future facts or events taking place in different but possible circumstances. Their usual form is *What would happen next?* or *What would happen if?...* Thus, Q2 questions are backward-causal questions seeking antecedents, whereas Q3 questions are forward-causal questions, which seek consequences.

This taxonomy for ISQ has proven to be suitable for classifying large amounts of student science questions (Ishiwa, Sanjosé, & Otero, 2012; Torres et al., 2012).

Hypotheses

In summary, previous research showed that: a) integrating linguistic and visual information is better for comprehension than providing only linguistic information to the students; b) watching dynamic images diminish the students' cognitive load when they try to comprehend time-dependent phenomena, and highly realistic images improves learning outcomes; c) handling devices in the laboratory could produce differences in comprehension with respect to only watching the devices; d) the higher the previous knowledge, the more suitable the

questions asked. From this theoretical approach, we formulated the following hypotheses:

- Hypothesis 1 (H1). Students in the Text condition will ask more association questions than students in the DVD condition or in the LAB condition.
- Hypothesis 2 (H2). Students in the three experimental conditions will generate large amounts of explanation questions, but the DVD or the LAB conditions could generate more explanation questions than the Text condition.
- Hypothesis 3 (H3). Students in the Text condition will ask less prediction questions than students in the LAB condition or in the DVD condition.
- Hypothesis 4 (H4) Students with high scientific knowledge will generate a higher proportion of questions including scientific concepts, laws and principles, than students with low scientific knowledge.

Some support for these hypotheses can be found in specific previous work on student science comprehension. Costa, Caldeira, Gallástegui and Otero (2000) has found that students focus on causality when they read expository texts so they ask significant amounts of explanation questions. Since causality is the basis of scientific information, substantial amounts of explanation questions are expected in the experimen-

tal conditions considered in the present study. Costa and colleagues (opus cit.) also obtained few prediction questions when students read texts about scientific phenomena, which agrees with other studies (Graesser & Bertus, 1998; Millis & Graesser, 1994). Therefore, we expect students in our reading condition to ask very few Q3 questions. In contrast, the possibility of watching the real operation of the devices in the DVD and the LAB conditions should make the mental representation of objects and events easy, and this could free cognitive resources for use in other comprehension processes, such as conjectures for causal factors:

- Study 1 was conducted to contrast hypotheses H1, H2, and H3 in a quantitative way.
- Study 2 was conducted to contrast hypothesis H4, and also two new hypotheses derived from some results of experiment 1. These new hypotheses are:

Hypothesis 5 (H5). In the LAB condition, the proportion of association questions will be similar to the one obtained in the DVD condition.

Hypothesis 6 (H6). In the DVD condition, the proportion of *What would happen if...* questions will be higher than the one obtained in the LAB condition, but the proportion of *Why* questions in the DVD condition will be lower than the one obtained in the LAB condition.

Method

Participants

Two samples, the first made up of 35 male and female Spanish 10th-grade secondary students (15-16 years old) and the second made up of 69 male and female Colombian 11th-grade secondary students (16-17 years old) participated in Experiment 1. They belonged to intact groups from secondary schools of big cities in the respective countries. This was the first time participants studied Physics & Chemistry, so they had poor scientific knowledge. The three groups of Colombian students were randomly assigned to the experimental conditions. One of the Spanish groups was assigned to the reading condition and the other group was randomly split up to assign students to the DVD or the LAB condition. Nine (7 Spanish and 2 Colombian) students in the LAB condition decided to give up on the experiment so complete data was obtained for 28 Spanish and 67 Colombian students.

Fifty-five 10th-grade (15-16 year olds) and 47 12th-grade (17-18 year olds) male and female Spanish secondary students participated in Experiment 2. Students belonged to four intact groups, two from 10th grade and two from 12th grade, from two educational centres in a big city, and all of them chose Physics and Chemistry as an elective in both academic levels.

The 12th-grade students had studied at least two physics courses, so they were supposed to have higher science knowledge than the 10th-grade students. In each academic level, the groups were randomly assigned to one of the two experimental conditions.

In both experiments the groups of participants did not explicitly present discriminating factors in any educational center, so they were considered to be equivalent in academic variables.

Materials

The devices used in both experiments were the following:

1. A double-cone that moves on a slope formed by two rigid bars (see description in Chart 1).
2. A *cartesian diver* made of a plastic pen cap and a piece of modelling clay float in a plastic bottle almost full of water and sealed with a cap. Squeezing the bottle by hand, the Cartesian diver sinks. But when one stops squeezing the bottle, the diver comes up again.

In the Text condition (considered only in experiment 1), students were given a booklet containing instructions followed by two short texts (with 91 and 92 words, respectively, in Spanish) describing the devices. Texts were designed to have three parts: (1) activation of a well known explanatory schema, (2) presenta-

tion of an event which is expected according to this schema, (3) introduction of an unexpected “dis-

crepant event” to produce perplexity. Chart 1 shows the double cone text.

Chart 1

The Double Cone Text in the Reading Condition

The climbing acrobat

We know that a round object rolls down an inclined plane. It does not move up unless it is motorized or pushed upwards. (*Activation of a well known explanatory schema*).

The *climbing acrobat* consists of a double cone (two cones joined at the base) and an inclined plane formed by two straight bars used as guides. If we place the bars parallel to each other and put the double cone on them, the object rolls down, as would be expected. (*The expected event agreeing with the explanatory schema*).

Nevertheless, when the angle of the bars is increased so that they form a *V-shape*, the double cone rolls up the slope without help! (*The discrepant event generating perplexity*).

The order of the two texts was counterbalanced in the booklets. Below each text there was a blank space for writing questions.

In the watching condition (studies 1 and 2), the operation of the devices was recorded on a DVD. One of the authors handled the devices and provided verbal information at the same time. The recorded verbal information was exactly the same as that which was included in the two texts. The video recording was made by an audio-visual technology expert using a professional camera. Sev-

eral *shots* were made from different perspectives when it was required to facilitate the visualization of the events. The film lasted 9:45 min.

In the LAB condition (both experiments), the two devices were physically placed in the school science laboratory.

Procedure

Throughout the experiment, an ethical protocol was respected and anonymity was protected. Both experiments were presented as edu-

cational tasks offering the opportunity to learn, but participation was voluntary.

Students were encouraged to ask the questions they might need in order to understand devices operation. To trigger student questions, we followed a clear method. First, the devices were selected due to their operating in a surprising, unexpected way so they could stimulate question asking according to Dillon (1990). Second, we tried to avoid well-known obstacles to question generation: a) social barriers were eliminated as students did not ask their questions aloud in front of other students; b) question asking was awarded with better grades according to their performance in the task proposed.

In order to avoid spurious questions (i.e. questions not necessary for comprehension), the researchers provide students with a particular task mental model (Llorens & Cerdán, 2012), different from *the more the questions asked the better the task*. Therefore: a) participants were instructed to understand the operation of the devices because they were going to be assessed in a second session; b) in the second session each participant would explain the devices to another (non-participant) student and the quality of the explanation would be assessed; c) answers to questions would be given before the second session so participants would give

better explanations to their partners.

At the end of the data collection sessions students were debriefed and informed that there would not be a second session.

In the Text condition (considered only in study 1) one of the researchers delivered the booklets to all the group in the classroom, the instructions were read aloud, students' doubts about the procedure were clarified, and the subjects were allowed to read the texts, re-read them, and write their questions down at their own pace in the space provided for that. This session lasted approximately 35 min: 15 min devoted to the organization and the reading of the instructions, and at least 20 min for reading and writing the questions.

The procedure in the DVD condition was similar in studies 1 and 2: group administration in the classroom and students had to write their questions down on a blank sheet. Students read the instructions and watched the DVD twice on a big screen (about 20 min). The second time around, the order of the devices shown was inverted. They wrote their questions down at their own pace during and after watching the DVD. In this condition, most participants took close to 35 minutes to complete the task.

In the LAB condition of both experiments, administration was individual and took place in the laboratory. The laboratory sessions were recorded with the permission

of adults and students. The instructions were read aloud to every participant. After clarifying doubts the first device was made to operate by the researcher. At the same time, he also provided verbal information to the participant, which was the same as that contained in the texts used in study 1. Next, the student was encouraged to handle the device at his/her own pace. The researcher made the second device operate following the same procedure. The order of operation of the devices was counterbalanced. The explanation and the instructions took about 10 min and the manipulation of the devices took typically 10-20 min per subject. Only one student needed more than 20 minutes to perform the task.

In the laboratory, students could ask their questions aloud whenever they required so, but the researcher could not answer them. If answers were provided, each one could add new information to the student's previous knowledge, so increasing it during the experiment. In that case, the different conditions considered in this study would not be comparable. We explained the operation of the devices to the students and answered all their questions off-line, at the end of the experimental sessions.

Measurements

The dependent variables were the amount of questions asked of every type: association ques-

tions (Q1), explanation questions (Q2) and prediction questions (Q3). Whereas the greater or lesser amount of questions asked depends on many non-controlled factors (for example, an extra motivation provided by the teachers; the need to gain academic credit in Science, etc.), the distribution of the questions asked in Q1, Q2 or Q3 type is expected to be related to the mental model the student is trying to build. Therefore, the proportions of each type of question asked were also considered. These proportions were obtained dividing the amount of each type of questions by the total amount of questions asked by each participant.

In order to know to what extent students focused their efforts of understanding on the operation of the devices, we studied the proportion of questions of any type addressed to the *target* information (Q_{target}), i.e. the amazing operation of the devices (the 'rolling upwards' in the double cone and the sinking and rising in the Cartesian diver).

Questions asked to clarify the instructions or the procedure were classified as procedural questions. Procedural questions represented a percentage about 1 per cent of the questions asked and they were discarded because our interest focused on the ISQ addressed to understand the scientific phenomena and not on the handling procedures.

In study 2 we took into account the number of questions involv-

ing scientific concepts, principles or laws (QSci) in each academic level. These questions should correspond to the obstacles the subject detected when he/she was trying to build the scientific mental representation. We considered any scientific term, including usual terms such as *speed* or *pressure*, when they were embedded in meaningful sentences in this context.

In study 1, two experts (one of them not belonging to the research team) independently categorized the total amount of questions asked. Some difficulties appeared to discriminate the backward causal questions (Q2) from the forward prediction questions (Q3) involving causal factors. Therefore, a new criterion was used: we decided to classify as Q3 the questions which mention in an explicit way new elements or concrete changes in the devices' setting. For example, the question "Is the liquid in the bottle necessarily water or could it be milk or oil?" was classified as Q3 because it mentioned new elements or changes not present in the devices, such as *milk* and *oil*. Thus, this question was interpreted as equivalent to "What would happen if the liquid in the bottle was milk or oil?", because the new elements suggested the student was attempting to go beyond the observed operation of the device. However, the question "The liquid in the bottle, does it have to be water?" was classified as Q2 because it did not men-

tion any new element or explicit changes. In this case the meaning of the question was considered to be equivalent to "Why does the liquid in the bottle have to be water?"

This new criterion allowed a substantial inter-coder agreement, resulting in a *kappa* index of .74. Additional disagreements were easily solved.

In experiment 2 we also obtained an inter-coder substantial agreement (*kappa* = .79). Disagreements were mainly mistakes and were solved by discussion.

Results

Study 1: Questions formulated under conditions of reading, watching or handling experimental devices

A total amount of 537 questions were obtained, corresponding to a global average of 5.65 questions per student. Table 1 shows the main data in experiment 1.

The total amount of questions followed a normal distribution, but the amounts of Q1, Q2, and Q3, did not. When the corresponding proportions were considered, PropQ1, PropQ2, PropQ3, and PropQtarget, only the proportion of Q3 was significantly different from a normal distribution, according to the Kolmogorov-Smirnov test. Thus, the proportions of Q1, Q2, Q3 and Qtarget were

Table 1

Mean Values per Subject (and Standard Deviations) for the Total Amount of Questions Asked and the Proportions of Each Type of Question in the Three Experimental Conditions

Exp.Condition	Qtot	PropQ1	PropQ2	PropQ3	PropQtarget
Reading (Text)	4.90 (2.15)	.57 (.25)	.41 (.25)	.02 (.07)	.34 (.28)
Watching (DVD)	7.41 (2.60)	.16 (.17)	.51 (.25)	.33 (.26)	.60 (.23)
Manipulating (LAB)	4.23 (2.35)	.15 (.32)	.73 (.38)	.12 (.21)	.59 (.32)

considered in further analyses. When PropQ3 was implied we used the non-parametric Mann-Whitney test to contrast our hypotheses; otherwise ANOVAs were used.

Simple ANOVAs for the proportions of Q1, Q2, and Qtarget taking the sub-sample (Spain/Colombia) as a between-subject factor did not show any significant effect (PropQ1 and PropQtarget: $F < 1$; PropQ2: $F = 1.005$; $p > .30$).

The same result was obtained performing the Mann-Whitney test for the Proportion of Q3 questions ($Z = -1.700$; $p = .089$). Therefore, the Colombian and the Spanish participants were considered together onwards.

Different analyses were performed taking the experimental condition (Text/ DVD/ LAB) as the between-subjects factor. Table 2 summarizes the main data from these analyses.

Table 2

Main Data from the Statistical Analyses Related to the Experimental Condition Factor. Post-hoc Pair Comparisons Addressed to Contrast the Hypotheses

	Global analyses	Post-hoc pair comparisons
PropQ1	$F(2,92) = 36.199$; $p < .001$ $h^2 = .44$; $p = 1.0$	Text > DVD; $p < .001$; Text > LAB; $p < .001$ (Dunnet) DVD \approx LAB; $p = .996$ (Sheffé)
PropQ2	$F(2,92) = 7.476$; $p = .001$ $h^2 = .14$; $p = .94$	Text < LAB; $p < .001$; Text \approx DVD; $p = .178$ (Dunnet) DVD \leq LAB; $p = .054$ (Sheffé)
PropQ3	$X^2 = 39.891$; $p < .001$ (Kruskal-Wallis)	Text < DVD; $p < .001$; LAB < DVD; $p = .008$ (M-Whit) Text \approx LAB; $p = .07$ (Mann-Whitney)
PropQtarget	$F(2,92) = 10.751$; $p < .001$ $h^2 = .19$; $p = .99$	Text < DVD; $p < .001$; Text < LAB; $p = .017$; (Sheffé) DVD \approx LAB; $p = .985$ (Sheffé)

In summary, the Text condition stimulated a higher proportion of association questions compared to the other two experimental conditions and hypothesis H1 was supported by the data; the DVD condition stimulated a higher proportion of prediction questions compared to the other two conditions, and H3 was partially supported; and the LAB condition seemed to stimulate a higher proportion of explanation questions with respect to the other two experimental conditions, although these proportions were high in the three experimental conditions, partially supporting H2. Concerning PropQtarget, the reading condition stimulated less proportion of these questions whereas the other two conditions, DVD and LAB obtained equivalent proportions of questions addressed to the “target” information. The last two results were contrasted again in experiment 2.

— Study 2: Replication and effect of prior knowledge on questions

about experimental devices in watching and handling conditions.

— Study 2 aimed to contrast H4 concerning the effect of previous knowledge on the questions asked, and also hypotheses H5 and H6 to replicate some interesting results obtained in study 1.

A total amount of 661 questions asked by the students in the two academic levels and in the two experimental conditions (DVD/LAB) was obtained. Table 3 shows the main data in the 10th and 12th grades. The variable Qsci accounted for questions including at least one scientific term (concept, law or principle) in their verbal formulation.

Most of the questions asked by 10th-grade students were explanation questions (3.9 questions/subject) with amounts clearly lower for association and prediction questions (about 0.8 questions/subject in both cases), whereas 12th-grade students were able to generate a

Table 3

Mean Values per Subject (and Standard Deviations) in Each Academic Level for the Total Amount of Questions Asked, the Proportion of Each Type of Question and the Amount of Questions Including Scientific Terms

Acad.Level	Qtot	PropQ1	PropQ2	PropQ3	Qsci
10 th -grade	5.45 (2.76)	.13 (.18)	.74 (.22)	.13 (.17)	1.35 (1.24)
12 th -grade	7.68 (4.83)	.19 (.21)	.57 (.22)	.25 (.23)	1.77 (2.22)

significant amount of prediction questions (1.9 questions/subject) in addition to the explanation questions (4.6 questions/subject).

More questions containing scientific terms were asked by 12th-grade students compared to 10th-grade students, but PropQsci, was higher for the 10th-grade students than for the 12th-grade students. The Mann-Whitney test

showed that these differences in PropQsci were not significant ($U = 1219.500$; $p = .620$), against the hypothesis H4.

Second, we contrasted hypotheses H5 and H6. Table 4 shows the mean values for the proportion of every type of question asked by students in the experimental conditions considered.

Table 4

Mean Values per Subject (and Standard Deviations) for the Total Amount of Questions Asked by Students and the Proportions of Each Type of Question in the Two Experimental Conditions

Exp. Condition	Qtot	PropQ1	PropQ2	PropQ3
Watching (DVD)	5.58 (3.14)	.14 (.21)	.61 (.26)	.25 (.23)
Manipulating (LAB)	7.53 (4.62)	.17 (.17)	.72 (.20)	.11 (.14)

None of the variables were normally distributed, except the proportion of Explanation questions, PropQ2. Thus, non-parametric Mann-Whitney tests for PropQ1 and PropQ3, and a simple ANOVA

for PropQ2 were performed. Table 5 shows the main data from the contrasts taking the experimental condition (DVD/ LAB) as the between-subjects factor.

Table 5

Main Data from the Statistical Analyses Addressed to Contrast the Hypotheses

Statistical parameters		
PropQ1	$U = 1099.000$; $p = .171$	DVD \approx LAB
PropQ2	$U = 878.000$; $p = .004$	DVD < LAB
PropQ3	$F(1,100) = 5.576$; $p = .020$; $h^2 = .05$; $p = .65$	DVD > LAB

Therefore the LAB condition stimulated a significantly higher proportion of explanation questions, but a significantly lower proportion of prediction questions than the DVD condition, supporting H6. Both conditions stimulated similar proportions of association questions, in agreement with H5.

Concerning the effect of previous knowledge, H4 was not supported by the data: students having high scientific knowledge did not use more scientific terms in their questions than students having low scientific knowledge.

Discussion

This research has shown that question formulation can be stimulated under particular conditions. Moreover, the format with which information is delivered to students can modify the kind of questions being generated, thereby giving evidence about the particular contribution of the scholar laboratory activities to science learning. The main points of this study can be summarized as follows:

1. Significant differences between the reading condition and the other experimental conditions (watching and handling) in the proportion of association questions were obtained, with a large effect size and high statistical power. Consistently, there

are significant differences between the reading condition and the handling condition in the proportion of other questions. These other questions are those addressing backward and forward causality. In the handling and watching conditions, the possibility of watching the devices in a realistic format seems to free enough cognitive resources to generate not only *Why* questions, but also *What would happen if...?* questions, and to focus on important information more than in the reading condition.

2. In experiment 1 the proportions of explanation questions were high in the three conditions, but the handling condition seemed to stimulate a higher proportion of explanation questions than the other two conditions. In addition the watching condition stimulated a higher proportion of prediction questions than the other two conditions. These results were supported by the data in study 2 for the watching and the handling conditions, although the statistical power was not high enough. Thus, even though the results have been replicated in Studies 1 and 2, a new, specific and controlled experiment is needed to increase their reliability. We can advance an explanation of these results. The possibility of handling the devices, and not only watching their operations,

could transform some prediction questions into explanation questions so generating the observed differences between the watching and the handling conditions. The substantial proportions of prediction questions in the form *What would happen if...?* obtained in the DVD condition suggest that some students tried to skip their comprehension obstacles by searching for causal factors and making conjectures. Although the conjectures could not be checked immediately, they represent an active effort to understand the devices mentally moving along causal chains. In the LAB condition, when an obstacle associated with a failed predictive inference is found, subjects can follow a *guess-and-check* strategy to immediately know what is really going to happen in the devices. Thus, instead of making a *What would happen if* question, a *Why* question would be generated as a consequence of the observed results of the handling done. This new conjecture has to be contrasted in a specific, controlled experiment.

3. It was expected that the higher the previous knowledge, the richer the attempted mental representation would be. In fact, the 12th-grade students generated a higher amount of questions than the 10th-grade students, especially prediction

questions. This result suggests that attempting *What would happen if* inferences requires certain level of previous knowledge involved in some causal scheme.

However, although the 12th-grade students obtained greater mean values than the 10th-grade students for questions containing scientific terms, the difference was not significant. That is an unexpected result since we expected the higher level students to recover and use their scientific knowledge more than the lower level students.

Educational consequences

The benefits of experimental activities in science education have been widely researched (Hofstein & Lunetta, 2004; Woolnough & Allsop, 1985). Learning tasks in the classroom should bring students as close as possible to scientists' work (Chinn & Malhotra, 2002). 'Good questions' are the origin of scientific knowledge so science education should stimulate question generation in students. Although asking questions produces significant benefits (Rosenshine, Meister, & Chapman 1996), students ask very few questions in the classroom (Dillon, 1988; Good, Slavings, Harel, & Emerson, 1987; Graesser & Persson, 1994).

Science texts are difficult to understand, even for college or university students, due to the complex mental representations readers have to build from common-usage language (Echevarria, 2006). Reading a text about scientific devices without any help from images implies a cognitive load for representing entities, and an additional difficulty in focusing on causal antecedents or consequences. Realistic images help students to focus on causal relationships, which is a goal of teachers. Since causality is essential in science comprehension, this is an interesting result giving psychological support to the use of experimental devices in science education. In addition, when students have the possibility of handling experimental devices, they can go beyond simply asking *why* questions and explore causal relationships by themselves using hypothetic-deductive reasoning, so generating predictive *What would happen if...* questions. Predictive inferences seemed to have helped these students to advance through causal chains using hypothetic-deductive reasoning. This is a specific contribution of experimental activities in laboratory conditions, which could be taken into account by teachers.

A higher knowledge was associated with the use of more hypothetic-deductive reasoning in our students. However, even at the end of the secondary education, stu-

dents seemed to have serious difficulties representing reality using science. This outcome alerts us about students' ability to apply what has been learned in the classroom to real problems supporting the need for more experimental activities which link reality to scientific knowledge.

Limitations of this study

Some limitations that have to be considered in present study. First, external validity is not guaranteed so results should not be straightforward transferred to other samples. Secondly, there were some procedure differences among the experimental conditions that could introduce some non-controlled effects: a) in the LAB condition, students asked their questions orally instead of writing them down on paper; b) in the LAB condition, student actions and questions were recorded whereas there was no recording device in the other two conditions. Thirdly, some unknown factors affected the total amount of questions asked by the students in some groups. Using proportions as the dependent variables, we tried to avoid these differences.

Since we replicated in study 2 the main results obtained in experiment 1, we expect that these limitations will not seriously affect the reliability of our conclusions, at least not in a qualitative manner.

References

- Barsalou, L. W. (2012). The human conceptual system. In M. J. Spivey, K. Mcrae, & M. F. Joannisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp 239-258). Cambridge: Cambridge University Press.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332. doi:10.1207/s1532690xcil0804_2.
- Chin, C. (2002). Student-generated questions: Encouraging inquisitive minds in learning science. *Teaching and Learning*, 23(1), 59-67.
- Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39. doi:10.1080/03057260701828101.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in Schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218. doi: 10.1002/sci.10001.
- Costa, J., Caldeira, M. H., Gallástegui, J. R., & Otero, J. (2000). An analysis of question asking on scientific texts explaining natural phenomena. *Journal of Research in Science Teaching*, 37, 602-614. doi: 10.1002/1098-2736(200008)37:6<602::AID-TEA6>3.0.CO;2-N.
- Dillon, J. T. (1988). The remedial status of student questioning. *Journal of Curriculum Studies*, 20, 197-210. doi:10.1080/0022027880200301.
- Dillon, J. T. (1990). *The practice of questioning*. New York: Routledge.
- Echevarría, M. A. (2006). ¿Enseñar a leer en la Universidad? Una intervención para mejorar la comprensión de textos complejos al comienzo de la educación superior. *Revista de Psicodidáctica*, 11(2), 169-188.
- Ferguson, E. L., & Hegarty, M. (1995). Learning with real machines or diagrams: application of knowledge to real-world problems. *Cognition and Instruction*, 13(1), 129-160. doi:10.1207/s1532690xcil301_4.
- Flammer, A. (1981). Towards a theory of question asking. *Psychological Research*, 43, 407-420. doi: 10.1007/bf00309225.
- Gangoso, Z. (2004). Un modelo para el proceso de resolución de problemas en Física. *Actas II Jornadas de Investigación en Resolución de Problemas en Física*. Córdoba (Argentina).
- García-Ros, R., & Pérez-González, F. (2011). Validez predictiva e incremental de las habilidades de autorregulación sobre el éxito académico en la universidad. *Revista de Psicodidáctica*, 16(2), 231-250.
- Good, T. L., Slavings, R. L., Harel, K. H., & Emerson, M. (1987). Students' passivity: A study of question asking in K-12 classrooms. *Sociology of Education*, 60, 181-199. doi: 10.2307/2112275.
- Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*, 2(3), 247-269. doi:10.1207/s1532799xssr0203_4.

- Graesser, A. C., & Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology, 95*, 524-536. doi: 10.1037/0022-0663.95.3.524.
- Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal, 31*, 104-137. doi:10.3102/00028312031001104.
- Graesser, A. C., & Zwaan, R. A. (1995). Inference generation and the construction of situation models. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Strategies and processing revisited. Essays in honor of Walter Kintsch* (pp. 117-139). Hillsdale, NJ: Erlbaum.
- Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr, & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert Simon* (pp. 285-318). Lawrence Erlbaum Associates, Hillsdale, NJ.
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722-738. doi: 10.1016/j.learninstruc.2007.09.013.
- Hofstein, A., & Lunetta, V. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28-54. doi: 10.1002/sce.10106.
- Ishiwa, K., Sanjosé, V., & Otero, J. (2012). Questioning and reading goals: Information-seeking questions asked on scientific texts read under different task conditions. Accepted for publication in the *British Journal of Educational Psychology*. doi:10.1111/j.2044-8279.2012.02079.x
- Llorens, A. C., & Cerdán, R. (2012). Assessing the comprehension of questions in task-oriented reading. *Revista de Psicodidáctica, 17*(2), 233-252. doi: 10.1387/rev.psicodidact.4496.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: Cambridge University Press.
- Millis, K., & Graesser, A. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language, 33*, 583-599. doi: 10.1006/jmla.1994.1028.
- Miyake, N., & Norman, D. A. (1979). To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18*, 357-364. doi: 10.1016/S0022-5371(79)90200-7.
- Otero, J. (2009). Question generation and anomaly detection in texts. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 47-59). New York: Routledge.
- Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction, 19*, 143-175. doi: 10.1207/S1532690XCI1902_01.
- Otero, J., Ishiwa, K., & Sanjosé, V. (2008). *Readers' questioning: Some hints for automated question generation*. Workshop on the Question Generation Shared Task and Evalu-

- ation Challenge. September 25-26, NSF, Arlington, VA.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66, 181-221. doi:10.3102/00346543066002181.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge: Cambridge University Press.
- Torres, T., Duque, J. K., Ishiwa, K., Sánchez, G., Solaz-Portolés, J. J., & Sanjose, V. (2012). Preguntas de los estudiantes de educación secundaria ante dispositivos experimentales. *Enseñanza de las Ciencias* 30(1), 49-60.
- Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, 21, 255-287. doi:10.1080/01638539609544959.
- Van der Meij, H. (1990). Question asking: To know that you do not know is not enough. *Journal of Educational Psychology*, 82, 505-512. doi:10.1037/0022-0663.82.3.505.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249-294. doi:10.3102/00346543063003249.
- Watts, M., Gould, G., & Alsop, S. (1997). Questions of understanding: categorising pupils, questions in science. *School Science Review*, 79(286), 57-63.
- Woolnough, B. E., & Allsop, T. (1985). *Practical work in science*. Cambridge: Cambridge University Press.

Vicente Sanjosé López is PhD in Physics, and professor of Science Education at the University of Valencia (Spain). His research focalizes on the psychological foundations of science learning and teaching. He is author of several research papers on comprehension of science texts, comprehension monitoring and problem-solving.

Tarcilo Torres Valois is a doctoral student in the Science Education program at the University of Valencia (Spain). He is lecturer at the University of Antioquia (Colombia).

Carlos Arturo Soto Lombana is PhD in Science Education and professor at the University of Antioquia (Colombia) where he was Dean of the Faculty of Education. He is the co-ordinator of the ‘Science and Mathematics Education Group’ and he is author of several research papers on the philosophical foundations of science education and on learning in interactive museums.

Received date: 20-02-2012

Review date: 05-09-2012

Accepted date: 10-12-2012

Efectos del formato en que se presenta la información científica sobre la autorregulación de los procesos de comprensión: la generación de preguntas

Vicente Sanjosé*, Tarcilo Torres**, y Carlos Soto**

*University of Valencia (Spain), **University of Antioquia (Medellín, Colombia)

Resumen

Formular preguntas es una de las acciones asociadas con procesos metacognitivos de regulación durante tareas de comprensión: los sujetos pueden generar preguntas destinadas a obtener información para salvar obstáculos de comprensión. Se realizaron dos experimentos para estimular, clasificar y analizar las preguntas de los estudiantes en tres condiciones diferentes: leer sobre el funcionamiento de dispositivos científicos experimentales, observarlos en un DVD o manipularlos en el laboratorio. Las preguntas fueron clasificadas usando una taxonomía sencilla. Teniendo en cuenta los principios multimedia de aprendizaje, las ventajas de imágenes realistas animadas para el aprendizaje de procesos dependientes del tiempo y el efecto de la actividad procesual-motora, se predijo diferente distribución de preguntas en las tres condiciones consideradas. Los resultados confirmaron lo esperado: la lectura de textos estimuló más las preguntas destinadas a describir las entidades, mientras que las condiciones de observar y manipular estimularon más las preguntas causales. Además, se analizó el efecto del conocimiento previo sobre las preguntas que incluyen conceptos científicos.

Palabras clave: Didáctica de las ciencias experimentales, autorregulación, generación de preguntas, dispositivos experimentales.

Abstract

Generating questions is a regulatory action associated with self-monitoring processes in comprehension tasks: subjects can ask 'information seeking questions' to solve comprehension obstacles. A sequence of two related experiments were conducted to trigger, classify and analyse questions asked under different conditions: reading a text about experimental scientific devices operating, watching these devices in a DVD and manipulating them in the LAB. Students' information seeking questions were classified using a simple taxonomy. Taking into account the multimedia learning principles, the advantages of realistic animations for understanding time-depending processes and the effect of the procedural-motor activity, we expected students would ask different questions under each of the aforementioned conditions. Results confirmed the expectations: the reading condition triggered more questions addressed to describe the entities while the watching and manipulation conditions stimulated more causal questions. In addition, the effect of prior knowledge on questions including scientific concepts was analyzed.

Keywords: Science education, self-regulation, question generation, experimental devices.

Agradecimientos: Investigación financiada por la Universidad de Valencia (Programa de Movilidad del Profesorado) y la Univesidad de Antioquía (Programa de Sostenibilidad 2011-2012).

Correspondencia: Vicente Sanjosé, Didáctica de las Ciencias Experimentales y Sociales, Universidad de Valencia. Avda Tarongers 4, 46022-Valencia. E-mail: vicente.sanjosé@uv.es

Introducción

Los profesores están de acuerdo acerca de la importancia de las preguntas de los estudiantes para un aprendizaje profundo (Chin, 2002; Chin y Osborne, 2008; Dillon, 1988; Flammer, 1981; Watts, Gould, y Alsop, 1997). En general, las actividades de regulación son buenos predictores del éxito académico (García-Ros y Pérez-González, 2011; Wang, Haertel, y Walberg, 1993) y hacer una pregunta es una actividad de auto-regulación.

En el contexto académico los tipos de preguntas más importantes son las dirigidas a resolver la falta de conocimiento, es decir, las preguntas que buscan información (ISQ son sus siglas en inglés). Nuestro objetivo es estudiar cómo el formato en que se presenta la información influye en la generación de ISQ. En concreto, queremos saber si la generación de preguntas varía cuando los estudiantes leen acerca de dispositivos experimentales científicos, los observan en un video o los manipulan físicamente. De acuerdo con los principios del aprendizaje multimedia (Mayer, 2001, 2005), el formato en que la información es presentada puede afectar las representaciones mentales que los estudiantes intentan construir para comprender la información (es decir, cómo funcionan los dispositivos). Así, el modelo *obstáculo-meta* sobre generación de preguntas (Otero, 2009) predice que diferentes tipos de preguntas serán formuladas cuando el

funcionamiento de los dispositivos científicos se presente a los sujetos en diferentes formatos: lectura de un texto sobre el funcionamiento de dispositivos científicos (en adelante condición Texto), visualización de los mismos dispositivos mientras están funcionando (condición DVD), o manipulación de esos dispositivos en el laboratorio (condición LAB o condición Laboratorio).

Las preguntas y el formato en que se presenta la información

En comparación con la información presentada solamente en formato texto, integrar información visual procedente de imágenes, gráficos y fotogramas con el texto puede mejorar la comprensión (Schnotz, 2005). Además, según la teoría de la carga cognitiva (Chandler y Sweller, 1991), visualizar un proceso o un procedimiento puede generar una reducción de la carga cognitiva en comparación con una situación en la que el proceso o el procedimiento tiene que ser reconstruido a partir de una serie de imágenes estáticas (por oposición a dinámicas), o leyendo un texto; entonces las animaciones (imágenes dinámicas) podrían ayudar a comprender los fenómenos científicos más que las imágenes estáticas. Höffler y Leutner (2007) encontraron una ventaja de tamaño medio de las animaciones instruccionales sobre las imágenes estáticas y tamaños del efecto sustancialmente mejores cuando la animación fue altamente realista.

Además, el manejo de dispositivos podría producir diferencias en los resultados de aprendizaje con respecto a la lectura o visualización sobre ellos. Algunos enfoques cognitivos recientes afirman que el sistema conceptual humano se fundamenta en sistemas modales específicos del cerebro, en la corporeidad (*embodiment*) y en el medio ambiente (Barsalou, 2012). Las representaciones conceptuales son representaciones modales, es decir, representaciones de los conceptos que dependen fuertemente de la modalidad de percepción (visual, táctil, etc.) Así, algunas diferencias en la comprensión deben aparecer cuando los estudiantes pueden manejar dispositivos del mundo real, o solamente pueden visualizarlos, o sólo pueden leer sobre ellos. Ferguson y Hegarty (1995) encontraron algunas ventajas de la manipulación de dispositivos reales con respecto a la visualización de imágenes o diagramas, en tareas de resolución de problemas por transferencia.

Las preguntas y el conocimiento previo

La influencia del conocimiento previo sobre las preguntas generadas ha sido investigada (Otero y Grasser, 2001; van der Meij, 1990) desde dos hipótesis: (a) falta de conocimiento del sujeto; (b) conflicto cognitivo. De acuerdo con la primera, una pregunta es generada cuando el sujeto no conoce la información particular que necesita saber. Sin embargo Miyake y Norman (1979) mostraron que la

falta de conocimiento no puede explicar la diversidad de las preguntas generadas. Graesser y Olde (2003) estudiaron las preguntas formuladas acerca de un dispositivo dañado por los participantes con diferentes niveles de experiencia y encontraron que cuanto más profunda es la comprensión tanto mayor será el número de preguntas adecuadas para reparar el aparato. Estos resultados corroboran la hipótesis del conflicto cognitivo: las preguntas formuladas dependen de las incoherencias o incompatibilidades entre la información suministrada y los conocimientos previos del sujeto.

Además, la naturaleza de las preguntas dependerá del conocimiento particular del sujeto. Cuando un estudiante tiene un conocimiento científico bajo sobre, por ejemplo, las tres leyes de Newton de la dinámica, es difícil para él formular una pregunta como «¿Cuál es la fuerza resultante que causa esta aceleración hacia la parte superior de la pendiente?». Es más probable que formule una pregunta como «¿Por qué se mueve el objeto hacia la parte superior de la pendiente?». La segunda formulación se refiere al conocimiento común y se asocia con la construcción de la representación mental o *modelo de situación*, mientras que la primera se asocia con el *modelo científico*.

Preguntas e inferencias en la comprensión de la ciencia

De acuerdo con el modelo *obstáculo-meta* (Otero et al., 2001), una

pregunta es generada cuando un sujeto encuentra un obstáculo de comprensión en el camino hacia la meta deseada, es decir, en la elaboración de cierta representación mental capaz de integrar la información proporcionada. La comprensión de la ciencia implica la construcción de representaciones mentales de alto nivel (Gangoso, 2004; Greeno, 1989) siendo la activación del conocimiento previo y la generación de inferencias procesos cognitivos importantes en la construcción de esas representaciones mentales. La creación de inferencias puede ser la actividad cognitiva más importante para la comprensión de la ciencia (Graesser y Zwaan, 1995). Las ISQ generadas en tareas de comprensión de la ciencia deben estar asociadas a los obstáculos específicos encontrados por los estudiantes: las inferencias intentadas, pero fallidas (Otero, Ishiwa, y Sanjosé, 2008).

Trabasso y Magliano (1996) identifican tres grandes categorías de inferencias generadas en el proceso de la comprensión consciente de texto narrativo: asociativas, explicativas y predictivas. Las *asociaciones* pretenden proporcionar información sobre las características, propiedades, relaciones y detalles descriptivos de las entidades (actores, objetos, eventos, actores humanos, acciones, etc.). Las *explicaciones* pretenden proporcionar razones acerca de por qué las cosas son tal como son. Por último, las *predicciones* están orientadas hacia el futuro e incluyen consecuencias de las ac-

ciones o acontecimientos y permiten anticipar eventos.

Esta taxonomía para las inferencias ha sido extendida al ámbito de los textos expositivos sobre ciencia y la correspondiente clasificación de preguntas ha sido propuesta por Otero et al. (2008). En primer lugar, las preguntas Q1 están vinculadas a *asociaciones* fallidas, frecuentemente se formulan usando *quién, qué, cómo, cuándo y dónde* y se dirigen a la necesidad de representar adecuadamente las entidades del sistema de que se trate, así como sus propiedades. Las preguntas Q2, o preguntas causales se formulan con frecuencia usando *por qué* y están vinculadas a explicaciones fallidas. Las preguntas predictivas, Q3 se generan a partir de inferencias fallidas hipotético-deductivas del tipo *si-entonces* y están dirigidas a prever hechos o eventos que podrían tener lugar en diferentes circunstancias. Su forma habitual es «¿qué pasaría después?» o «¿qué pasaría si ...?». Por tanto, las preguntas Q2 son preguntas de antecedente causal, mientras que las preguntas Q3 son preguntas de consecuente causal.

Esta taxonomía para las ISQ ha demostrado ser adecuada para clasificar grandes cantidades de preguntas formuladas por estudiantes de ciencias (Ishiwa, Sanjosé, y Otero, 2012; Torres et al., 2012).

Hipótesis

En resumen, la investigación previa muestra que: a) la integra-

ción de la información lingüística y visual es mejor para la comprensión que proporcionar solamente información lingüística a los estudiantes; b) la visualización de imágenes dinámicas disminuye la carga cognitiva de los alumnos cuando tratan de comprender los fenómenos que se extienden en el tiempo y las imágenes altamente realistas mejora los resultados de aprendizaje; c) la manipulación de dispositivos en el laboratorio podría producir diferencias en la comprensión con respecto a la mera visualización de los dispositivos; d) cuanto mayor es el conocimiento previo, más apropiadas para la comprensión son las preguntas formuladas. A partir de este marco teórico, se formulan las siguientes hipótesis:

- Hipótesis 1 (H1). Los estudiantes en la condición Texto formularán más preguntas asociativas que los estudiantes en la condición DVD o en la condición LAB.
- Hipótesis 2 (H2). Los estudiantes en las tres condiciones experimentales generaran cantidades importantes de preguntas explicativas, pero en las condiciones DVD o LAB podrían generar más preguntas explicativas que en la condición Texto.
- Hipótesis 3 (H3). Los estudiantes en la condición Texto generarán menos preguntas predictivas que los estudiantes en la condición LAB o en la condición de DVD.
- Hipótesis 4 (H4). Los estudiantes con conocimiento científico alto

generarán una mayor proporción de preguntas que incluyan conceptos científicos, leyes y principios, que los estudiantes con conocimiento científico bajo.

Cierto apoyo para estas hipótesis se puede encontrar en trabajos previos, específicos sobre la comprensión de la ciencia. Costa, Caldeira, Gallástegui y Otero (2000) encontraron que los estudiantes se centran en la causalidad cuando leen textos expositivos por lo que formulan una cantidad importante de preguntas explicativas. Dado que la causalidad es la base de la información científica, cantidades sustanciales de preguntas explicativas son esperables en las condiciones experimentales consideradas en el presente estudio. Costa et al. (2008) también obtienen pocas preguntas predictivas cuando los estudiantes leen textos sobre fenómenos científicos, lo que concuerda con otros estudios (Graesser y Bertus, 1998; Millis y Graesser, 1994). Por lo tanto, se espera que los estudiantes de la condición Texto formulen muy pocas preguntas Q3. Sin embargo, la posibilidad de visualizar el funcionamiento real de los dispositivos en las condiciones DVD y LAB debe facilitar la creación de representaciones mentales de los objetos y eventos y esto podría liberar recursos cognitivos para su uso en los procesos de comprensión, tales como conjeturas para factores causales.

El estudio 1 se realizó para contrastar las hipótesis H1, H2 y H3. El estudio 2 se realizó para contrastar

la hipótesis H4 y también dos nuevas hipótesis derivadas de algunos resultados del estudio 1; estas nuevas hipótesis son las siguientes:

- Hipótesis 5 (H5). En la condición LAB, la proporción de preguntas asociativas será similar a las obtenidas en la condición DVD.
- Hipótesis 6 (H6). En la condición de DVD, la proporción de preguntas del tipo «¿qué pasaría si...?» será mayor que las obtenidas en la condición LAB, pero la proporción de preguntas del tipo «¿por qué?» en la condición de DVD será menor que las obtenidas en la condición de LAB.

Método

Participantes

En el estudio 1 se consideraron dos muestras inicialmente. La primera estaba formada por 35 estudiantes españoles de ambos sexos de décimo grado de Educación Secundaria (15-16 años). La segunda muestra la componían 69 estudiantes de ambos sexos colombianos de undécimo grado de Secundaria (16-17 años). Pertenecían a grupos intactos de centros educativos situados en grandes ciudades de los países respectivos. Era la primera vez que los participantes estudiaban un curso de Física y Química, así que tenían bajo conocimiento científico en el momento del estudio. Los tres gru-

pos de estudiantes colombianos fueron asignados al azar a las condiciones experimentales. Uno de los dos grupos españoles fue asignado a la condición de lectura y el otro grupo fue dividido aleatoriamente para asignar a los estudiantes a la condición DVD o a LAB. Nueve alumnos (siete españoles y dos colombianos) asignados a la condición LAB decidieron abandonar el experimento, por lo que se obtuvieron datos completos de 28 españoles y 67 estudiantes colombianos.

En el estudio 2 participaron cincuenta y cinco estudiantes de 10.º grado (15-16 años) y cuarenta y siete estudiantes de 12.º grado (17-18), españoles y de ambos sexos. Los estudiantes pertenecían a cuatro grupos intactos, dos de 10.º grado y dos de 12.º grado, de dos centros educativos de Secundaria en una ciudad importante. Todos ellos eligieron Física y Química como optativa en ambos niveles académicos. Los estudiantes de 12.º grado habían estudiado al menos dos cursos de Física, por lo que se supone que tenían un mayor conocimiento científico que los estudiantes de 10.º grado. En cada nivel académico, los grupos fueron asignados aleatoriamente a una de las dos condiciones experimentales.

En ambos experimentos, los grupos de participantes no presentaron explícitamente factores discriminantes en ninguno de los centros educativos, por lo que se consideraron equivalentes en las variables académicas.

Materiales

Los dispositivos utilizados en ambos experimentos fueron los siguientes:

1. Un doble cono que se mueve en una pendiente formada por dos barras rígidas (véase la descripción en el Cuadro 1).
2. Un *diablillo cartesiano* hecho a partir de una tapadera de bolígrafo de plástico y de un trozo de plastilina, que se deja flotar en una botella de plástico transparente casi llena de agua y cerrada con su tapón. Al presionar la botella con las manos, el diablillo cartesiano se hunde. Pero cuando se deja de presionar la botella, el diablillo asciende nuevamente.

Cuadro 1

El Texto del Doble Cono en la Condición Texto: «El acróbata trepador»

Sabemos que los objetos redondos ruedan hacia abajo por un plano inclinado a no ser que tengan motor o sean empujados. (*Activación de un esquema explicativo previo*).

El *acrobata trepador* consiste en un doble cono (dos conos unidos por su base) y un plano inclinado formado por dos guías rectas. Si colocamos las guías en paralelo y ubicamos el doble cono sobre ellas, éste rueda hacia abajo como se espera. (*El evento esperado de acuerdo al esquema explicativo*).

Sin embargo, cuando aumentamos el ángulo de las guías y éstas forman una V, el doble cono rueda hacia la parte alta de la rampla sin ayuda! (*Evento discrepante que provoca perplejidad*).

En la condición Texto (considerada solamente en el estudio 1), se repartió un cuadernillo a los estudiantes con las instrucciones seguidas de dos textos cortos (con 91 y 92 palabras, respectivamente, en español) que describen los dispositivos. Los textos fueron diseñados con tres partes: (1) la activación de un esquema explicativo previo; (2) la presentación de un evento conocido que se espera que suceda de acuerdo con este esquema; (3) la

introducción de un «evento discrepante», inesperado, para provocar perplejidad. El Cuadro 1 muestra la versión en español de un texto sobre el doble cono.

El orden de los dos textos fue contrabalanceado en los cuadernillos. Debajo de cada texto se disponía de un espacio en blanco para escribir las preguntas.

En la condición de visualización (estudios 1 y 2), el funcionamiento de los dispositivos fue grabado en

un DVD. Uno de los investigadores hizo funcionar los dispositivos mientras proporcionaba información verbal. La información verbal registrada fue exactamente la misma que la incluida en los dos textos. La grabación de vídeo fue hecha por un experto en tecnología audio-visual con una cámara profesional. Se grabaron varias tomas o planos desde diferentes perspectivas para facilitar la visualización del funcionamiento de los dispositivos. La película duró 9:45 minutos.

En la condición de LAB (en ambos experimentos), los dos dispositivos se colocaron físicamente en un lugar adecuado en el laboratorio escolar.

Procedimiento

A lo largo del experimento se respetó un protocolo ético y el anonimato fue preservado. Ambos experimentos fueron presentados como tareas educativas con la oportunidad de aprender, pero la participación fue voluntaria.

Los estudiantes fueron animados a formular las preguntas que necesitaran para comprender el funcionamiento de los dispositivos. Para provocar las preguntas de los estudiantes se siguió un procedimiento muy determinado. En primer lugar, los dispositivos fueron seleccionados debido a su funcionamiento sorprendente e inesperado, por lo que podían estimular las preguntas, de acuerdo con Dillon (1990). En segundo lugar, se trató de evitar los

obstáculos conocidos para la generación de preguntas: a) las barreras sociales fueron eliminadas debido a que los estudiantes no formulaban preguntas en voz alta delante de otros estudiantes, b) la formulación de preguntas fue premiada con un pequeño incremento en la calificación, de acuerdo a su desempeño en la tarea propuesta.

Para evitar preguntas espurias (es decir, preguntas no necesarias para la comprensión), los investigadores trataron de proporcionar a los estudiantes un modelo mental de trabajo (Llorens y Cerdán, 2012), diferente del modelo mental *cuantas más preguntas formuladas mejor*. Para ello: a) los participantes fueron instruidos para comprender el funcionamiento de los dispositivos ya que iban a ser evaluados en una segunda sesión; b) en esa segunda sesión cada participante explicaría los dispositivos a otro estudiante (no participante) y la calidad de la explicación sería evaluada; c) se entregarían a los estudiantes las respuestas a sus preguntas antes de la segunda sesión, para que los participantes dieran mejores explicaciones.

Al final de la primera sesión se informó a los estudiantes que no se realizaría una segunda sesión.

En la condición Texto (considerada solamente en el estudio 1) la administración de los cuadernillos se llevó a cabo en el aula al grupo completo. Uno de los investigadores leyó las instrucciones en voz alta y aclaró las dudas de los estudiantes sobre el procedimiento. Tras ello los

sujetos leyeron los textos, los releeron y escribieron sus preguntas en el espacio dispuesto para ello a su propio ritmo. Esta sesión duró aproximadamente 35 minutos: 15 minutos dedicados a la organización y lectura de las instrucciones, y aproximadamente 20 minutos a leer y escribir las preguntas.

El procedimiento en la condición DVD fue similar: la administración fue grupal en el aula y los estudiantes tenían que escribir sus preguntas en una hoja en blanco. Los estudiantes visualizaron el DVD dos veces en una pantalla grande (aproximadamente 20 minutos). La segunda vez, el orden de los dispositivos mostrados fue invertido. Ellos escribieron sus preguntas a su propio ritmo durante y después de la visualización del DVD. En esta condición, la mayoría de los participantes empleó cerca de 35 minutos para completar la tarea.

En la condición LAB la administración fue individual y tuvo lugar en el laboratorio. Las sesiones de laboratorio se grabaron en formato DVD con el permiso de adultos y de estudiantes. Uno de los investigadores explicó el experimento repartió las instrucciones y las leyó en voz alta con cada estudiante. Después de aclarar las dudas, el primer dispositivo fue puesto en funcionamiento por uno de los investigadores. Al mismo tiempo, también se proporcionó información verbal a cada estudiante que era la misma que la contenida en los textos utilizados en la condición Texto. A continuación

se animó a cada alumno a manipular el dispositivo por sí mismo. El investigador puso en funcionamiento el segundo dispositivo siguiendo el mismo procedimiento. El orden en el funcionamiento de los dispositivos fue contrabalanceado. La explicación y las instrucciones supusieron alrededor de 10 minutos y la manipulación de los dispositivos entre 10-20 minutos por sujeto. Solamente un estudiante necesitó más de 20 minutos para realizar la tarea.

En la condición de laboratorio, los estudiantes formularon sus preguntas en voz alta, pero el investigador no las respondió. Si las respuestas hubieran sido proporcionadas, cada una podría añadir nueva información a los conocimientos previos de los estudiantes, que se hubiera incrementado durante el experimento. En ese caso, las diferentes condiciones consideradas en este estudio podrían no ser comparables. En un momento posterior al experimento, al final de las sesiones experimentales, los investigadores explicaron el funcionamiento de los dispositivos a los estudiantes y respondieron a todas sus preguntas.

Medidas

La variable dependiente fue la cantidad de preguntas de cada tipo: preguntas asociativas (Q1), preguntas explicativas (Q2) y preguntas predictivas (Q3). La mayor o menor cantidad de preguntas formuladas por un estudiante depende de muchos factores no controlados como,

por ejemplo, una motivación extra creada por los profesores, la necesidad de obtener una bonificación en ciencias, etc. La distribución de las preguntas Q1, Q2 o Q3 formuladas se debe relacionar con el modelo mental que el estudiante está tratando de construir por lo que se consideraron las proporciones de cada tipo de pregunta; estas proporciones fueron obtenidas dividiendo la cantidad de cada tipo de preguntas por el número total de preguntas de cada participante.

Con el fin de conocer en qué medida los alumnos centraron sus esfuerzos en la comprensión del funcionamiento de los dispositivos, se estudió la proporción de preguntas de cada tipo dirigida a la información *diana* (Q_{target}), es decir, el funcionamiento sorprendente de los dispositivos (el rodar «hacia arriba» en el doble cono y el «hundimiento y ascenso» en el diablillo cartesiano).

Las preguntas formuladas para aclarar las instrucciones o el procedimiento fueron clasificadas como preguntas procedimentales y representaron un porcentaje cercano al 1% de las preguntas formuladas. Este tipo de preguntas fue descartado porque nuestro interés se centró en las ISQ dirigidas a comprender los fenómenos científicos y no en las dirigidas a los procedimientos de manipulación.

En el Experimento 2 se tuvo en cuenta el número de preguntas que involucran conceptos científicos, principios o leyes (Q_{Sci}) en cada nivel académico. Estas preguntas deberían corresponder a los obstáculos

los detectados por los sujetos en su intento de construir una representación mental científica. Se consideró cualquier término científico, incluyendo términos usuales tales como *velocidad* o *presión* siempre que se incluyeron en frases con sentido en este contexto.

En el estudio 1, dos expertos (uno de ellos no perteneciente al equipo de investigación) clasificaron de forma independiente el número total de preguntas formuladas. Algunas dificultades aparecieron a la hora de diferenciar las preguntas causales hacia atrás (Q2) de las preguntas de predicción (Q3) de tipo hipotético-deductivo involucrando factores causales. La discusión posterior no solucionó estas ambigüedades. Por tanto, un nuevo criterio fue usado: se decidió clasificar como Q3 solamente las preguntas que mencionaran nuevos elementos o cambios concretos en la configuración de los dispositivos de una manera explícita. Por ejemplo, la pregunta «¿el líquido en la botella es necesariamente agua o puede ser leche o aceite?» fue clasificada como Q3 porque menciona elementos nuevos o modificaciones que no están presentes en los dispositivos, tales como *leche* y *aceite*. Por ello esta pregunta fue interpretada como equivalente a «¿qué pasaría si el líquido de la botella fuera leche o aceite? porque los nuevos elementos sugieren que el estudiante está tratando de ir más allá del funcionamiento observado del dispositivo. Sin embargo, la pregunta «¿el líquido en la botella, tiene que ser

agua?» fue clasificada como Q2 ya que no menciona ningún elemento nuevo o cambios explícitos. En este caso el significado de la pregunta fue considerado como equivalente a la de «¿por qué el líquido en la botella tiene que ser agua?». Este nuevo criterio provocó un incremento del acuerdo inter-jueces, resultando un índice *kappa* de 0.74 (acuerdo substancial). Algunos desacuerdos posteriores se resolvieron con facilidad.

Resultados

Estudio 1: Preguntas formuladas en condiciones de lectura, visualización o manipulación de dispositivos experimentales

Se obtuvo un total de 537 preguntas, que corresponde a un promedio global de 5.65 preguntas por alumno. La Tabla 1 muestra los principales datos del estudio 1.

Tabla 1

Valores Medios por Sujeto (y Desviaciones Estándar) en el Número Total de Preguntas Formuladas y Proporciones de Cada Tipo de Pregunta en las Tres Condiciones Experimentales

Condición Exp.	Qtot	PropQ1	PropQ2	PropQ3	PropQtarjet
Lectura (Text)	4.90 (2.15)	.57 (.25)	.41 (.25)	.02 (.07)	.34 (.28)
Visionado (DVD)	7.41 (2.60)	.16 (.17)	.51 (.25)	.33 (.26)	.60 (.23)
Manipulación (LAB)	4.23 (2.35)	.15 (.32)	.73 (.38)	.12 (.21)	.59 (.32)

La cantidad total de preguntas formuladas siguió una distribución normal, pero las cantidades de preguntas de tipo Q1, Q2 y Q3 no lo hicieron. Cuando las correspondientes proporciones (PropQ1, PropQ2, PropQ3 y PropQtarjet) fueron consideradas, solamente la proporción de Q3 fue significativamente diferente de una distribución normal, de acuerdo a la prueba de Kolmogorov-Smirnov. Por ello, en los análisis posteriores se utilizó la prueba no-paramétrica Mann-Whitney para contrastar las hipótesis sobre PropQ3 y en el resto de variables se utilizó el test ANOVA.

Los ANOVA simples para las proporciones de Q1, Q2 y Qtarjet,

tomando la submuestra (España/Colombia) como un factor entre-sujetos, no mostraron ningún efecto significativo (PropQ1 y PropQtarjet: $F < 1$; PropQ2: $F = 1.005$, $p > .30$). El mismo resultado se obtuvo realizando la prueba de Mann-Whitney para la proporción de preguntas Q3 ($Z = -1.700$, $p = 0.089$). Por lo tanto, los participantes colombianos y españoles fueron considerados como una sola muestra en adelante.

Los distintos análisis realizados tomando la condición experimental (Texto/DVD/LAB) como el factor entre sujetos se resumen en la Tabla 2.

Tabla 2

Principales Datos de los Análisis Estadísticos Relacionados con el Factor Condición Experimental. Comparaciones Post-hoc entre Pares Dirigidas a Contrastar las Hipótesis

	Análisis Global	Comparaciones <i>post-hoc</i> entre pares
PropQ1	$F(2,92) = 36.199; p < .001$ $h^2 = .44; P = 1.0$	Text > DVD; $p < .001$; Text > LAB; $p < .001$ (Dunnet) DVD \approx LAB; $p = .996$ (Sheffé)
PropQ2	$F(2,92) = 7.476; p = .001$ $h^2 = .14; P = .94$	Text < LAB; $p < .001$; Text \approx DVD; $p = .178$ (Dunnet) DVD \leq LAB; $p = .054$ (Sheffé)
PropQ3	$\chi^2 = 39.891; p < .001$ (Kruskal-Wallis)	Text < DVD; $p < .001$; LAB < DVD; $p = .008$ (M-Whit) Text \approx LAB; $p = .07$ (Mann-Whitney)
PropQtarget	$F(2,92) = 10.751; p < .001$ $h^2 = .19; p = .99$	Text < DVD; $p < .001$; Text < LAB; $p = .017$; (Sheffé) DVD \approx LAB; $p = .985$ (Sheffé)

La condición de Texto estimuló una mayor proporción de preguntas asociativas que las otras dos condiciones experimentales y, por tanto, la hipótesis H1 recibe apoyo de los datos. La condición DVD estimuló una mayor proporción de preguntas predictivas en comparación con las otras dos condiciones, con lo que la hipótesis H3 queda parcialmente apoyada. Finalmente, la condición LAB pareció estimular una mayor proporción de preguntas explicativas que las otras dos condiciones experimentales, aunque estas proporciones fueron altas en las tres condiciones experimentales; este resultado apoya parcialmente la hipótesis H2. En cuanto a las preguntas específicamente dirigidas a la información *diana*, la condición Texto estimuló menor proporción de preguntas de este tipo, mientras que las condiciones DVD y LAB produjeron una proporción equivalente de ellas. Estos dos últimos resultados fueron contrastados de nuevo en el estudio 2.

Estudio 2: Replicación y efecto del conocimiento previo sobre la formulación de preguntas acerca de dispositivos experimentales en condiciones de visualización y manipulación

El experimento 2 tuvo como objetivo contrastar la hipótesis H4 referente al efecto del conocimiento previo sobre las preguntas formuladas, así como las hipótesis H5 y H6 para replicar algunos resultados interesantes obtenidos en el estudio 1.

Se registró un total de 661 preguntas de los estudiantes en los dos niveles de enseñanza y en las dos condiciones experimentales consideradas (DVD/LAB). En la Tabla 3 se muestran los principales datos registrados en los grados 10.º y 12.º. La variable Qsci representa las preguntas que incluyen al menos un término científico (concepto, ley o principio) en su formulación verbal.

La mayoría de las preguntas formuladas por los estudiantes de

Tabla 3

Valores Medios por Cada Sujeto (y Desviaciones Estándar) en Cada Nivel Académico en la Cantidad Total de Preguntas, Proporción de Cada Tipo de Pregunta y Cantidad de Preguntas que Incluyen Términos Científicos

Nivel Acad.	Qtot	PropQ1	PropQ2	PropQ3	Qsci
10.º grado	5.45 (2.76)	.13 (.18)	.74 (.22)	.13 (.17)	1.35 (1.24)
12.º grado	7.68 (4.83)	.19 (.21)	.57 (.22)	.25 (.23)	1.77 (2.22)

10.º grado fueron explicativas (3.9 preguntas/sujeto) con cantidades claramente inferiores a las preguntas asociativas y predictivas (alrededor de 0.8 preguntas/sujeto en ambos casos), mientras que los estudiantes de 12.º grado fueron capaces de generar una cantidad significativa de preguntas predictivas (1.9 preguntas/sujeto), además de preguntas explicativas (4.6 preguntas/sujeto).

Los estudiantes de 12.º grado formularon un mayor número de pregun-

tas con inclusión de términos científicos que los estudiantes de 10.º grado, pero el número de PropQsci fue mayor entre los estudiantes de 10.º grado que entre los estudiantes de 12.º grado. En todo caso, la prueba de Mann-Whitney mostró que estas diferencias en PropQSci no son significativas ($U = 1219.500$, $p = 0.620$), en contra de la hipótesis H4.

En la Tabla 4 se exponen datos con respecto a las hipótesis H5 y H6: la proporción de cada tipo de

Tabla 4

Valores Medios por Sujeto (y Desviaciones Estándar) en el Total de Preguntas Formuladas y Proporciones de Cada Tipo de Pregunta en las Dos Condiciones Experimentales

Condición Exp.	Qtot	PropQ1	PropQ2	PropQ3
Visualización(DVD)	5.58 (3.14)	.14 (.21)	.61 (.26)	.25 (.23)
Manipulación(LAB)	7.53 (4.62)	.17 (.17)	.72 (.20)	.11 (.14)

pregunta en las condiciones experimentales consideradas.

Solamente la proporción de preguntas explicativas PropQ2 se distribuyó normalmente; en consecuencia, se utilizó la prueba no-paramétrica de Mann-Whitney para PropQ1 y PropQ3 y el test ANOVA simple

para PropQ2. En la Tabla 5 se muestran los principales datos de los contrastes que toman la condición experimental (DVD/LAB).

En la condición LAB se observa una proporción significativamente mayor de preguntas explicativas, pero una proporción significativa

Tabla 5

Principales Datos del Análisis Estadístico Dirigidos a Contrastar las Hipótesis 5 y 6

Parámetros estadísticos		
PropQ1	$U = 1099.000; p = .171$	DVD \approx LAB
PropQ2	$U = 878.000; p = .004$	DVD < LAB
PropQ3	$F(1,100) = 5.576; p = .020; h^2 = .05; p = .65$	DVD > LAB

mente menor de preguntas predictivas, que la condición DVD, confirmando la hipótesis H6. Ambas condiciones estimularon una proporción similar de preguntas asociativas, de acuerdo con H5.

Con respecto al efecto del conocimiento previo, la hipótesis 4 no se ve apoyada por los datos: los estudiantes con un conocimiento científico más alto no usaron más términos científicos en la formulación de sus preguntas que los estudiantes con un conocimiento científico más bajo.

Discusión

Esta investigación demuestra que la formulación de preguntas puede ser estimulada en ciertas condiciones; más aún, que el formato con la que la información se presenta a los alumnos afecta al tipo de preguntas que se generan, mostrando claramente la contribución particular al aprendizaje de las ciencias de las actividades experimentales en los laboratorios escolares. Los aspectos principales pueden resumirse del modo siguiente:

1. Se verifican diferencias significativas entre la condición de

lectura y las demás condiciones experimentales (visualización y manipulación) en la proporción de preguntas asociativas, con un tamaño del efecto grande y alta potencia estadística. Consistentemente, hay diferencias significativas entre la condición de la lectura y la condición de manipulación en la proporción del resto de tipos de preguntas. Estas preguntas afectan a la causalidad, dirigidas bien hacia los antecedentes bien hacia los consecuentes (hacia el futuro). En las condiciones de manipulación y visualización la posibilidad de observar los dispositivos en un formato realista parece liberar suficientes recursos cognitivos para generar no sólo preguntas del tipo «¿por qué...?» sino también del tipo «¿qué pasaría si...?» y concentrarse en la información importante, más que en la condición de lectura.

2. En el experimento 1 la proporción de preguntas explicativas es alta en las tres condiciones, pero la condición de manipulación parece estimular una mayor proporción de preguntas explicativas que las otras dos condiciones.

Además la condición de visualización estimula una proporción mayor de preguntas predictivas que las otras dos condiciones. Estos resultados fueron apoyados por los datos del estudio 2 en las condiciones de visualización y de manipulación, aunque no se alcanza significatividad estadística. Por ello, a pesar de que los resultados se han replicado en los experimentos 1 y 2, sería necesario llevar a cabo un nuevo experimento controlado, específico, para aumentar la fiabilidad de estos resultados.

Puede avanzarse una explicación a estos resultados. La posibilidad de manipular los dispositivos, y no sólo de visualizar su funcionamiento, podría transformar algunas de las preguntas predictivas en preguntas explicativas generando las diferencias observadas entre las condiciones de visualización y manipulación. La proporción de preguntas predictivas de la forma «¿qué pasaría si...?» obtenida en la condición DVD sugiere que algunos estudiantes tratan de salvar los obstáculos de comprensión por medio de la búsqueda de factores causales mediante conjeturas. A pesar de que esas conjeturas no pueden comprobarse inmediatamente en esta condición, representan un esfuerzo activo por comprender los dispositivos moviéndose mentalmente a lo largo de cadenas causales. En la condición LAB, cuando se detecta un obstáculo asociado a una inferencia predictiva, los suje-

tos pueden seguir una estrategia de *conjetura-y-verificación* para saber de inmediato lo que realmente pasa en los dispositivos. Por ello, en lugar de hacer una pregunta de tipo «¿qué pasaría si...?», se genera una pregunta de tipo «¿por qué?» como consecuencia de los resultados observados de la manipulación realizada. En todo caso, esta explicación tendría que ser contrastada en futuras investigaciones.

Se esperaba que cuanto mayor fuera el conocimiento previo, más rica fuera la representación mental que se intenta construir. De hecho, los estudiantes de 12.º grado generaron mayor cantidad de preguntas que los estudiantes de 10.º grado, especialmente de preguntas predictivas; este resultado sugiere que intentar realizar inferencias hipotético-deductivas requiere cierto nivel de conocimientos previos involucrados en algún esquema causal.

Sin embargo, aunque los estudiantes de 12.º grado obtuvieron mayores valores medios que los estudiantes de 10.º grado para las preguntas que contienen términos científicos, la diferencia no es estadísticamente significativa. Este es un resultado inesperado ya que se preveía que los estudiantes de nivel superior podrían recuperar y utilizar sus conocimientos científicos más que los alumnos de nivel inferior.

Consecuencias educativas

Los beneficios de las actividades experimentales en la enseñanza

de las ciencias han sido ampliamente investigados (Hofstein y Lunetta, 2004; Woolnough y Allsop, 1985). Las tareas de aprendizaje en el aula deben llevar a los estudiantes lo más cerca posible de la labor desarrollada por los científicos (Chinn y Malhotra, 2002). Las «buenas preguntas» son el origen del conocimiento científico, así que la educación científica debe estimular la generación de preguntas en los estudiantes. Aunque la formulación de preguntas produce beneficios significativos (Rosenshine, Meister, y Chapman, 1996), los estudiantes hacen muy pocas preguntas en el aula (Dillon, 1988; Good, Slavings, Harel, y Emerson, 1987; Graesser y Person, 1994).

Los textos de ciencia son difíciles de entender, incluso para los estudiantes universitarios, debido a las representaciones mentales complejas que los lectores tienen que construir a partir del lenguaje común (Echevarría, 2006). La lectura de un texto acerca de un dispositivo científico sin ninguna ayuda de imágenes implica una carga cognitiva para la representación de las entidades y una dificultad adicional para centrarse en antecedentes o consecuentes causales. Las imágenes realistas pueden ayudar a los estudiantes a centrarse en las relaciones de causalidad, que son los objetivos de los profesores. Dado que la causalidad es esencial en la comprensión de la ciencia, esto es un resultado interesante que da apoyo psicológico a la utilización de dispositivos experimentales en la educación científica. Además,

cuando los estudiantes tienen la posibilidad de manipular los dispositivos experimentales, pueden ir más allá de la simple formulación de la pregunta «¿por qué...?» y explorar relaciones causales por sí mismos usando razonamiento hipotético-deductivo, generando preguntas predictivas del tipo «¿qué pasaría si...?». Las inferencias predictivas parecen haber ayudado a los estudiantes a avanzar a través de las cadenas causales utilizando razonamiento hipotético-deductivo. Esta es una contribución específica de las actividades experimentales en condiciones de laboratorio, que podría ser tomada en consideración por los profesores.

Un conocimiento científico superior se asoció con el uso de más inferencias hipotético-deductivas en nuestros estudiantes. Sin embargo, incluso al final de la Educación Secundaria, los estudiantes parecen tener serias dificultades para representar la realidad mediante el uso de la ciencia. Este resultado alerta sobre la capacidad de los estudiantes para aplicar lo que han aprendido en el aula a los problemas reales. Parece necesario realizar más actividades experimentales que vinculen el conocimiento científico con la realidad.

Limitaciones de este estudio

Algunas limitaciones han de ser consideradas en el presente estudio. En primer lugar, la validez externa no está garantizada por lo que los resultados de los estudios 1 y 2 no deberían ser directamente transferidos

a otras muestras. En segundo lugar, hubo algunas diferencias de procedimiento entre las condiciones experimentales que podrían introducir algunos efectos no controlados: a) en la condición de LAB, los estudiantes formularon sus preguntas oralmente en lugar de escribirlas en un papel; b) en la condición de LAB las acciones de los estudiantes y sus preguntas fueron grabadas mientras que no hubo ningún dispositivo de grabación en las otras dos condiciones;

c) en tercer lugar, algunos factores desconocidos afectaron la cantidad total de preguntas formuladas por los alumnos de algunos grupos; utilizando las proporciones como variables dependientes tratamos de evitar los efectos de estas diferencias.

Ahora bien, dado que los principales resultados obtenidos en el estudio 1 se replicaron en el estudio 2, cabe confiar en que estas limitaciones no afecten cualitativamente a las conclusiones de esta investigación.

Referencias

- Barsalou, L.W. (2012). The human conceptual system. En M. J. Spivey, K. Mcrae, y M. F. Joannisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp 239-258). Cambridge: Cambridge University Press.
- Chandler, P., y Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332. doi:10.1207/s1532690xci0804_2.
- Chin, C. (2002). Student-generated questions: Encouraging inquisitive minds in learning science. *Teaching and Learning*, 23(1), 59-67.
- Chin, C., y Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39. doi:10.1080/03057260701828101.
- Chinn, C. A., y Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218. doi: 10.1002/sce.10001.
- Costa, J., Caldeira, M. H., Gallástegui, J. R., y Otero, J. (2000). An analysis of question asking on scientific texts explaining natural phenomena. *Journal of Research in Science Teaching*, 37, 602-614. doi: 10.1002/1098-2736(200008)37: 6<602::AID-TEA6>3.0.CO;2-N.
- Dillon, J. T. (1988). The remedial status of student questioning. *Journal of Curriculum Studies*, 20, 197-210. doi:10.1080/0022027880200301.
- Dillon, J. T. (1990). *The practice of questioning*. New York: Routledge.
- Echevarría, M. A. (2006). ¿Enseñar a leer en la Universidad? Una intervención para mejorar la comprensión de textos complejos al comienzo de la educación superior. *Revista de Psicodidáctica*, 11(2), 169-188.
- Ferguson, E. L., y Hegarty, M. (1995). Learning with real machines or diagrams: application of knowledge to real-world problems. *Cognition and Instruction*, 13(1), 129-160. doi:10.1207/s1532690xci1301_4.

- Flammer, A. (1981). Towards a theory of question asking. *Psychological Research*, 43, 407-420. doi: 10.1007/bf00309225.
- Gangoso, Z. (2004). Un modelo para el proceso de resolución de problemas en Física. *Actas II Jornadas de Investigación en Resolución de Problemas en Física*. Córdoba (Argentina).
- García-Ros, R., y Pérez-González, F. (2011). Validez predictiva e incremental de las habilidades de autorregulación sobre el éxito académico en la universidad. *Revista de Psicodidáctica*, 16(2), 231-250.
- Good, T. L., Slavings, R. L., Harel, K. H., y Emerson, M. (1987). Students' passivity: A study of question asking in K-12 classrooms. *Sociology of Education*, 60, 181-199. doi: 10.2307/2112275.
- Graesser, A. C., y Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*, 2(3), 247-269. doi:10.1207/s1532799xssr0203_4.
- Graesser, A. C., y Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95, 524-536. doi: 10.1037/0022-0663.95.3.524.
- Graesser, A. C., y Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, 31, 104-137. doi:10.3102/00028312031001104.
- Graesser, A. C., y Zwaan, R. A. (1995). Inference generation and the construction of situation models. En C. A. Weaver, S. Mannes, y C. R. Fletcher (Eds.), *Discourse comprehension: Strategies and processing revisited. Essays in honor of Walter Kintsch* (pp. 117-139). Hillsdale, NJ: Erlbaum.
- Greeno, J. G. (1989). Situations, mental models, and generative knowledge. En D. Klahr, y K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert Simon* (pp. 285-318). Lawrence Erlbaum Associates, Hillsdale, NJ.
- Höfler, T. N., y Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738. doi: 10.1016/j.learninstruc.2007.09.013.
- Hofstein, A., y Lunetta, V. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88(1), 28-54. doi: 10.1002/scs.10106.
- Ishiwa, K., Sanjosé, V., y Otero, J. (2012). Questioning and reading goals: Information-seeking questions asked on scientific texts read under different task conditions. Accepted for publication in the *British Journal of Educational Psychology*. doi:10.1111/j.2044-8279.2012.02079.x
- Llorens, A. C., y Cerdán, R. (2012). Assessing the comprehension of questions in task-oriented reading. *Revista de Psicodidáctica*, 17(2), 233-252. doi: 10.1387/rev.psicodidact.4496.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. En R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: Cambridge University Press.
- Millis, K., y Graesser, A. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language*, 33, 583-599. doi: 10.1006/jmla.1994.1028.
- Miyake, N., y Norman, D. A. (1979). To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 18, 357-364. doi: 10.1016/S0022-5371(79)90200-7.

- Otero, J. (2009). Question generation and anomaly detection in texts. En D. Hacker, J. Dunlosky, y A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 47-59). New York: Routledge.
- Otero, J., y Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction*, 19, 143-175. doi: 10.1207/S1532690XCI1902_01.
- Otero, J., Ishiwa, K., y Sanjosé, V. (2008). *Readers' questioning: Some hints for automated question generation*. Workshop on the question generation shared task and evaluation challenge. September 25-26, NSF, Arlington, VA.
- Rosenshine, B., Meister, C., y Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66, 181-221. doi:10.3102/00346543066002181.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. En R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge: Cambridge University Press.
- Torres, T., Duque, J. K., Ishiwa, K., Sánchez, G., Solaz-Portolés, J. J., y Sanjose, V. (2012). Preguntas de los estudiantes de educación secundaria ante dispositivos experimentales. *Enseñanza de las Ciencias*, 30(1), 49-60.
- Trabasso, T., y Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, 21, 255-287. doi:10.1080/01638539609544959.
- Van der Meij, H. (1990). Question asking: To know that you do not know is not enough. *Journal of Educational Psychology*, 82, 505-512. doi:10.1037/0022-0663.82.3.505.
- Wang, M. C., Haertel, G. D., y Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249-294. doi: 10.3102/00346543063003249.
- Watts, M., Gould, G., y Alsop, S. (1997). Questions of understanding: categorising pupils, questions in science. *School Science Review*, 79(286), 57-63.
- Woolnough, B. E., y Allsop, T. (1985). *Practical work in science*. Cambridge: Cambridge University Press.

Vicente Sanjosé López es doctor en física por la Universidad de Valencia y profesor titular de Didáctica de las Ciencias Experimentales en esa misma universidad. Su investigación atiende a las bases psicológicas del aprendizaje y la didáctica de las ciencias. Ha publicado diversos artículos sobre comprensión de textos científicos, control de la comprensión y resolución de problemas.

Tarcilo Torres Valois es estudiante de doctorado en el programa de investigación en didácticas específicas de la Universidad de Valencia. Es también profesor de física y didáctica de las ciencias en la Universidad de Antioquia, Colombia.

Carlos Arturo Soto Lombana es doctor en didáctica de las ciencias por la Universidad de Valencia y profesor titular en la Universidad de Antioquia, donde es decano en la facultad de Educación. Dirige el grupo de investigación «Educación en Ciencias Experimentales y Matemáticas» (GECEM) y es autor de varios artículos sobre bases filosóficas de la educación científica y sobre enseñanza en los museos interactivos de ciencias.

Fecha de recepción: 20-02-2012

Fecha de revisión: 05-09-2012

Fecha de aceptación: 10-12-2012

