Psicothema

Psicothema 2014, Vol. 26, No. 3, 385-390
doi: 10.7334/psicothema2013.252

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

# The effects of question-generation training on metacognitive knowledge, self regulation and learning approaches in Science

Francisco Cano García, Ángela García, A.B.G. Berbén, M.C. Pichardo and Fernando Justicia
Universidad de Granada

## Abstract

**Background:** Although much research has examined the impact of question generation on students' reading comprehension and learning from lectures, far less research has analysed its influence on how students learn and study science. The present study aims to bridge this knowledge gap. **Method:** Using a quasi-experimental design, three complete ninth-grade science classes, with a total of 72 students, were randomly assigned to three conditions (groups): (G1) questioning-training by providing prompts; (G2) question-generation without any explicit instruction; and (G3) no question control. Participants' pre-test and post-test self-reported measures of metacognitive knowledge, self-regulation and learning approaches were collected and data analysed with multivariate and univariate analyses of covariance. **Results:** (a) MANCOVA revealed a significant effect for group; (b) ANCOVAs showed the highest average gains for G1 and statistically significant between-group differences in the two components of metacognition: metacognitive knowledge and self-regulation; and (c) the direction of these differences seemed to vary in each of these components. **Conclusions:** Question-generation training influenced how students learned and studied, specifically their metacognition, and it had a medium to large effect size, which was somewhat related to the prompts used.

***Keywords:*** Questioning; learning approaches; self-regulation; question generation; metacognition.

## Resumen

***Efectos del entrenamiento en generación de preguntas sobre el conocimiento metacognitivo, la autorregulación y los enfoques de aprendizaje en Ciencias. Antecedentes:*** aunque muchas investigaciones han examinado el impacto de la generación de preguntas en la comprensión lectora de los estudiantes, pocas de ellas han analizado su influencia en cómo los estudiantes aprenden y estudian en Ciencias. Este estudio pretrende reducir ese déficit de conocimientos. **Método:** utilizando un diseño cuasi-experimental, tres clases de Ciencias de noveno grado (N = 72 estudiantes) fueron asignadas aleatoriamente a tres grupos: (G1) entrenamiento en generación de preguntas mediante indicaciones; (G2) generación de preguntas sin instrucción explícita; y (G3) grupo control. Los participantes proporcionaron medidas de autoinforme pre-test y post-test en conocimiento metacognitivo, autorregulación y enfoques de aprendizaje, las cuales fueron analizadas mediante análisis de covarianza univariados y multivariados. **Resultados:** a) el MANCOVA reveló un efecto significativo del factor Grupo; b) los ANCOVAs mostraron el cambio promedio más elevado para G1 y diferencias significativas en dos componentes de la metacognición (conocimiento y autorregulación); y c) la dirección de esas diferencias pareció variar en cada uno de esos componentes. **Conclusiones:** entrenar en generar preguntas influenció el modo de aprender de los estudiantes (e.g., metacognición), influencia de tamaño medio a grande relacionada con las indicaciones utilizadas.

***Palabras clave:*** hacer preguntas; enfoques de aprendizaje; autorregulación; generación de preguntas; metacognición.

Questioning is an essential feature of inquiry-based science and one of the key disciplinary practices prioritised in current science education reform by organisations such as the National Research Council (NRC; 2012). An ideal learner—self-regulated and active—is a 'person who asks deep questions and searches for answers to thought-provoking questions' (Otero & Graesser, 2001, pp. 143-144).

Questioning has a long and well-documented list of benefits for students' learning, ranging from increasing motivation to engendering productive discussion in the classroom and to directing knowledge construction (see Chin & Osborne, 2008; Yu, 2009). Despite its potential value, students' questioning in the classroom is, however, infrequent and mainly focused on low-level questions that involve minimal inferences and surface cognitive processing (Chin, Brown, & Bruce, 2002; Graesser & Person, 1994). Accordingly, much research has been directed at fostering question-generation as a means of improving students' reading comprehension (see Wong, 1985; Rosenshine, Meister, & Chapman, 1996) and learning from lectures (e.g., King, 1992). Less research, however, has studied the influence of teaching students to generate questions on some variables linked to their learning in science, such as their strategic self-regulation, metacognitive knowledge and learning approaches. The present paper aims to bridge this knowledge gap by analysing the possible effects on these variables of an intervention programme on question-generation.

Received: August 20, 2013 • Accepted: April 23, 2014
Corresponding author: Francisco Cano García
Facultad de Ciencias de la Educación
Universidad de Granada
18071 Granada (Spain)
e-mail: fcano@ugr.es

*Intervention studies on question-generation*

Questioning has been referred to in the literature as: (a) a comprehension-fostering cognitive strategy (e.g., Palincsar & Brown, 1984), (b) a metacognitive or comprehension-monitoring activity (e.g., Rosenshine et al., 1996; Wong, 1985), and (c) a learning strategy involved in meaningful generation of learning (Wittrock, 1990).

A first and sizeable group of studies has been focused on training students in question-generation during or after reading or listening to a passage, as a means of fostering their comprehension. The idea is that questioning engages students in an active and deeper processing (Craig & Lockhart, 1972), which improves their reading comprehension and retention of target content (King, 1994). However, this expectation has not always been supported by experimental investigations, due in part, perhaps, to the wide range of methods used to stimulate question-generation (for a review, see Wong, 1985; Rosenshine et al., 1996).

A second group of studies has taught students to generate questions as a means of improving their metacognitive, or comprehension-monitoring, activity. The idea is that, according to metacognitive theory, self-monitoring instruction would foster students' evaluation of their own reading comprehension and learning, leading to the enhancement of their reading and studying (for a review, see Wong, 1985).

A third group of studies (e.g., King, 1992, 1994) has been carried out in the theoretical framework of Wittrock's model of generative learning (see Lee, Lim, & Grabowski, 2007; Wittrock, 1990), in which students learn better by engaging in active knowledge construction. Within this framework, questioning is an effective learning strategy that facilitates such generative processing through different processes, among which developing metacognitive skills to monitor their understanding of information (metacognition) is vital (Lee et al., 2007). Although "the effectiveness of self-questioning is attributed to both its cognitive and metacognitive functions" (King, 1992, p. 305), the latter are critical for generative learning: "without metacognition, students can become overwhelmed in determining what information is relevant to their needs and what they need to do to refine known strategies" (Land, 2000, p. 73).

*Methods of instruction and criterion measures*

Researchers have used methods of instruction in question generation that range from regular instruction to modelling and reciprocal teaching (Rosenshine et al., 1996; Wong, 1985). Although in their review of 26 empirical studies, Rosenshine et al. (1996) did not find significant differences between these in terms of effect size, they suggest for future research a number of instructional elements to support student learning (e.g., the provision of cue cards listing the prompts for generating questions and the use of practice in small groups). In addition, they emphasised the importance of the different procedural prompts (e.g., signal words; generic question stems and generic questions) of which the most effective in terms of effect size were the generic question stems (e.g., "what is a new example of ... ?"; "how is it related to ...?"). Their effect was larger when students were provided with these prompts rather than practising question-generation without any explicit facilitation, and when experimenter-developed comprehension tests rather than standardised tests were used. Finally, they indicated that for future studies, it would be worth continuing to explore the value of prompts by including three treatments (conditions or groups): questioning-training providing prompts, question-generation without any explicit instruction and no-question control. In some recent studies (e.g., Bugg & Daniel, 2012; Weinstein, McDermott, & Roediger, 2010), students who generated and answered their own questions performed better (e.g., remembering the information in the texts) than did a group who re-read the texts. However, these studies did not include Rosenshine et al.'s (1996) three suggested treatments.

Concerning the criterion measures, in many of the 27 studies reviewed by Wong (1985) researchers used a single index of effectiveness, and this is the method she suggested for future studies incorporating multiple dependent variables (DVs). Some years later, King (1992) trained college students in a guided cooperative questioning strategy to learn from lectures, and assessed it through several scores, which improved after the intervention. Although this extension in both the number of DVs and the domain (from texts to lectures) seemed remarkable, it failed to comply with Rosenshine et al.'s (1996) recommendation of carrying out more research on the effectiveness of questioning in different content areas, including how students learn and study.

*Students' learning in Science*

Two research perspectives, learning approaches and self-regulation, are generally applied when explaining how students learn in science. The literature on these perspectives is quite extensive and beyond the scope of this paper, however, an overview of their core constructs may be illustrative where they are relevant to the purpose of the current investigation.

*Learning approaches*. These are one of the key constructs to understand students' learning processes during science lessons (e.g., Appleton & Beasley, 1994) and refer to the ways, deep or surface, in which students go about their learning (e.g., Baeten, Kyndt, Struyven, & Dochy, 2010; Entwistle & McCune, 2004). Learning approaches consist of motives and strategies elicited by different personal and situational factors (e.g., perceptions of the teaching context) and influence students' learning outcomes. The deeper the students' learning approaches, the higher the quality of their learning outcomes (e.g., Arias, Cabanach, Núñez, & González-Pienda, 1998; Baeten et al., 2010).

Two types of study have been undertaken to examine the link between questioning and learning approaches: quantitative and qualitative. In quantitative experimental studies, researchers have tried to induce a deep learning approach through different interventions such as including adjunct questions in text and questions after reading (see Marton & Säljö, 1997) and creating constructivist learning environments (see Baeten et al., 2010). Their results, however, indicated how difficult it is for students to change the way they learn to a deeper approach. By contrast, in qualitative studies researchers have observed the types of question students ask and related it to their learning approaches (e.g., Chin et al., 2002; Chin & Brown, 2000; Pedrosa de Jesus, Almeida, Teixeira-Dias, & Watts, 2006). Their results indicated that those students who adopted a surface approach tended to ask basic or surface questions (i.e., factual), whereas those who adopted a deep approach tended to ask wonderment or deep questions (i.e., comprehension). However, such results appear somewhat limited: Pedrosa de Jesús et al. (2006) used ten participants and Chin et al.

(2002) observed only six target students, who represented extreme learning approaches, rather than the usual larger samples. Thus, there is very little empirical or experimental evidence linking training students to generate questions to an improvement in their learning approaches.

*Self-regulation*. This is a key element of the new developments in cognitive information-processing and constructivist learning theories (Shell et al., 2005). Although there are a considerable diversity of theoretical and methodological perspectives and conceptions (e.g., Zimmerman, 2001, 2008) on self-regulation, all tend to agree that it is a 'proactive process that students use to acquire academic skills, such as setting goals, selecting and deploying strategies and self-monitoring one's effectiveness' (Zimmerman, 2008, p. 166). Knowledge-building, self-regulated strategies and question-asking in class (high and low level) are different aspects of students' perceptions of their own strategic self-regulation, mapped by Shell et al. (2005), who developed the Student Perceptions of Classroom Knowledge-Building Scale (SPOCK) to assess these behaviours within a specific course. The SPOCK assesses the regulation of cognition components (e.g., planning), through its self-regulated strategies sub-scale, but not the knowledge component (e.g., procedural knowledge). However, students who are good strategy-users coordinate the two components (Pressley, Borkowski, & Schneider, 1987). Therefore, this minor limitation of the SPOCK should be addresed.

Some researchers (e.g., Núñez et al., 2011) have reported highly satisfactory results of intervention programmess focused on training studying and self-regulation strategies. Participants were, however, university students, and self-questioning was apparently embedded in the set of training strategies. Although good strategy users actively construct knowledge by engaging in a great variety of generative learning strategies (e.g., question-asking, self-regulation) (e.g., Pintrich, 2004; Pressley et al., 1987; Shell et al., 2005), to the best of our knowledge, very little research has studied how training secondary students to generate questions affects their strategic self-regulation and metacognitive knowledge in science classes. By using these criterion measures in addition to those proposed in the framework of learning approaches, the differences between the three treatments or groups suggested by Rosenshine et al. (1996) would be discovered.

In the light of the above limitations, the present study has two aims: a) to assess the efficacy of the question-generation programme in terms of learning approaches, strategic self-regulation and metacognitive knowledge, and b) to determine the value of providing prompts for questioning. Despite the scarcity of research in this area, following the review of literature our expectations were:

1. Students trained in question-generation would have the highest mean score for metacognitive knowledge and strategic self-regulation (Chin & Osborne, 2008; King, 1992, 1994; Landa, 2000; Lee et al., 2007; Rosenshine et al., 1996; Wong, 1985).
2. Assuming the value of prompts, the means of the scores in Rosenshine et al.'s (1996) proposed groups would be ordered in the direction of: questioning training providing prompts > question-generation without any explicit instruction > no question control (Bugg & McDaniel, 2012; King, 1992, 1994; Rosenshine et al., 1996; Weinstein et al., 2010).

## Method

### Participants

Participants were 72 ninth-grade students (35 boys, 37 girls) at a secondary school, in an urban district in Granada (Andalusia, Spain), who came from all social strata. They were all Caucasian, had a mean age of 14.44 years (*SD* = .69), and were enrolled in one of three separate classes of the same science course, each class being taught by a different teacher.

### Design

A pre-test, post-test quasi-experimental design was adopted for the study. The above-mentioned classes were randomly assigned to one of the three conditions or groups mentioned: G1 (questioning training providing prompts) ($n$ = 29); G2 (question-generation without any explicit instruction ($n$ = 17); and G3 (no question control) ($n$ = 26).

### Measures

*Approaches to learning*. These were assessed using a modified version of the revised two-factor version of the Learning Process Questionnaire (R-LPQ-2F, Kember, Biggs, & Leung, 2004). This questionnaire, which was slightly modified to specifically assess learning approaches within a science class, included 22 items, grouped into four subscales: surface motive, surface strategy, deep motive, and deep strategy (e.g., *I try to relate new science material, as I am reading it, to what I already know on that topic*), corresponding to the two learning approach dimensions, Deep and Surface, proposed by its authors. Students gave responses on a Likert-type scale, from 1 (*never or rarely true of me*) to 5 (*always or almost always true of me*). The inducted-sample reliability coefficients (i.e., the reliability coefficient reported by the instrument developers) were .82 for Deep approach, and .71 for Surface approach. The current-sample internal consistency coefficients (Cronbach's alphas) were .83 for Deep approach, and .58 for Surface approach, the latter being lower than desirable but still within the acceptable range for measures developed and used for research purposes (Nunnally, 1978).

*Strategic self-regulation*. This was assessed through the SPOCK (Shell et al., 2005), using four subscales containing 24 items focused on students' perceptions of their own strategic self-regulation: Self-regulated strategy use (e.g., *In this class, I set goals for myself which I try to accomplish*); Knowledge building (e.g., *As I study topics in other classes, I try to think about how they relate to the topics I am studying in this class*); High-level question-asking (e.g., *In this class, I ask questions to help me better understand the things I am trying to learn*) and Low-level question-asking (e.g., *In this class, I ask questions to help me prepare for tests*). Responses were gathered on a 5-point Likert-type scale ranging from 1 (almost never) to 5 (almost always). The inducted-sample Cronbach's alpha values of these subscales were as follows: .81, .84, .92 and .91. The current-sample Cronbach's alpha values of these subscales were as follows: .79, .79, .93 and .81.

*Metacognitive knowledge*. This was measured using the 9-item Metacognitive knowledge sub-scale from the "Junior Metacognitive Awareness Inventory (Jr. MAI) … (which includes) … declarative, procedural, and conditional knowledge of cognition" (Sperling, Howard, Miller, & Murphy, 2002, p. 55). Participants rated these items (e.g., *I try to use strategies that have worked in the past*) on a 5-point Likert scale ranging from 1 (never) to 5 (always) and their scores were reliable (the inducted-sample Cronbach's alpha = .82; the current-sample Cronbach's alpha = .99).

*Procedure*

The teacher of each class met the investigator separately for one session (of 1 hour) prior to the beginning of the study for training in the procedure to follow in his/her corresponding group. Likewise, the students from these groups (G1, G2 and G3) were trained by their teachers during four 20-minute sessions. Then, in the last ten minutes of their final-term science classes (240' in total), students had to use what they had learnt during training.

Students in G1 received a sheet including several generic question stems and examples. They were informed about the importance of questioning in science learning: its role and value in general and especially when the questions are deep or thought-provoking. They were then taught to generate different types of question: simple questions (e.g., who …?); intermediate questions checking their understanding (e.g., What does X mean?) and linking together two ideas from the lesson (e.g., how is X similar to Y?); and complex questions eliciting deep-reasoning patterns (e.g., why did an event occur?). Following a guided cooperative questioning strategy, they used these question stems "to generate their own specific questions on the material being studied. Then in small groups or pairs they put their questions to each other and answered each other's questions" (King, 1994, p. 340). Students in G2 received the same information. However, they were untrained in questioning and not provided with the question stems, but directed to ask and answer each other's questions. Students in G3 also worked in pairs but in a similar way to one of the groups used in Davey and McBride's (1986) study, they thought of the meaning of some terms used during the lesson and looked up the dictionary definitions instead of generating or answering questions.

*Data analyses*

In the present study, data analyses were conducted in two phases. First, a series of preliminary analyses, which examined the descriptive statistics of scores in all study variables and the normality of their distribution, were conducted. Second, two types of analysis, which took into account pre-test score differences, were used (a) to evaluate the overall intervention effects (a multivariate analysis of covariance, MANCOVA) and (b) to test for significant intervention effects on each of the post-test scores (one-way analyses of covariance, ANCOVAs).

## Results

*Preliminary analyses*

Table 1 shows the descriptive information for the three groups in pre-test and post-test scores.

Examination of the distribution of these variables indicated that they were all considered appropriate for use in parametric statistical analyses since their scores did not significantly depart from normality according to D'agostino and Pearson's K2 (1973) test and the Lagrange multiplier test of Jarque and Bera (1987).

The overall effect of the intervention was determined by using a multivariate analysis of covariance, with the Group (condition) as within-subject factor, after adjusting for pre-test score differences. Results showed a significant multivariate effect for Group (Wilks' λ (14,116) = .667, $p$ = .038). Next, a series ANCOVAs was conducted individually on each DV, statistically controlling for any differences on pre-scores.

*ANCOVAs*

Statistical assumptions underlying ANCOVAs (Ato & Vallejo, 2007) were met: (a) slopes equal to zero (the interaction between Group and covariate was statistically significant for each DV); (b) homogeneity of regression slope (the relationship between the DV and the covariate was the same in each treatment group); and c) homogeneity of variance (homoscedasticity) (Levene's $F$ tests were not significant, $p$>.05). The results of the ANCOVAs with equal slopes and pre-test scores as covariates are summarised in Table 2.

*Table 1*

Means and standard deviations for pre-test and post-test DVs scores in three conditions (Groups)

| Groups | G1 | | G2 | | G3 | |
|---|---|---|---|---|---|---|
| DVs | M | SD | M | SD | M | SD |
| Pretest | | | | | | |
| Knowledge-building | 3.01 | 0.72 | 2.77 | 0.72 | 2.78 | 0.65 |
| Self-regulated strategy use | 3.50 | 0.75 | 3.65 | 0.64 | 3.53 | 0.66 |
| High-Level Question-Asking | 3.49 | 0.90 | 3.02 | 0.90 | 3.11 | 1.03 |
| Low-Level Question-Asking | 3.16 | 0.89 | 3.26 | 1.05 | 3.23 | 0.92 |
| Deep approach | 3.02 | 0.61 | 2.97 | 0.47 | 2.86 | 0.60 |
| Surface approach | 3.36 | 0.48 | 3.56 | 0.45 | 3.46 | 0.57 |
| Metacognitive knowledge | 4.06 | 0.44 | 4.09 | 0.50 | 4.15 | 0.49 |
| Post-test | | | | | | |
| Knowledge-building | 3.02 | 0.81 | 2.73 | 0.69 | 2.50 | 0.73 |
| Self-regulated strategy use | 3.60 | 0.78 | 3.30 | 0.64 | 3.15 | 0.68 |
| High-Level Question-Asking | 3.29 | 0.90 | 2.73 | 0.67 | 2.72 | 0.78 |
| Low-Level Question-Asking | 3.52 | 1.05 | 2.99 | 0.85 | 3.12 | 0.92 |
| Deep approach | 2.96 | 0.75 | 2.78 | 0.62 | 2.58 | 0.63 |
| Surface approach | 3.35 | 0.49 | 3.54 | 0.45 | 3.50 | 0.45 |
| Metacognitive knowledge | 4.04 | 0.49 | 3.50 | 0.75 | 3.90 | 0.50 |

Note: G1 = questioning training providing prompts); G2 = question-generation without any explicit instruction; and G3 = no question control

*Table 2*

Results of the ANCOVAs: adjusted means (i.e., controlling for pre-test score differences), F-values, significance levels and effect size for the post-test scores of the different DVs

| | G1 | G2 | G3 | | | |
|---|---|---|---|---|---|---|
| DVs | M | M | M | F | p | $\eta^2$ |
| Knowledge-building | 2.94 | 2.80 | 2.55 | 2.46 | .093 | .06 |
| Self-regulated strategy use | 3.63 | 3.24 | 3.17 | 4.90 | .010 | .12 |
| High-Level Question-Asking | 3.17 | 2.85 | 2.79 | 2.50 | .089 | .07 |
| Low-Level Question-Asking | 3.54 | 2.97 | 3.12 | 2.56 | .085 | .07 |
| Deep approach | 2.92 | 2.77 | 2.64 | 1.06 | .203 | .04 |
| Surface approach | 3.40 | 3.47 | 3.49 | .357 | .701 | .01 |
| Metacognitive knowledge | 4.06 | 3.51 | 3.87 | 7.06 | .002 | .17 |

Note: G1 = questioning training providing prompts); G2 = question-generation without any explicit instruction; and G3 = no question control; df = 2, 70.

An initial glance at these means indicated that in all the DVs except Surface approach, students in G1 (questioning training providing prompts) had the highest scores. A second glance, based on test statistic (F) and p-values revealed that (a) some between-groups differences in question-asking in class (high and low level) emerged, but did not reach the level of significance, and (b) statistically significant differences were focused on two DVs: Self-regulation strategies ($F = 4.90$, $p<.01$, $\eta = .12$) and Metacognitive knowledge ($F= 7.06$, $p<.01$, $\eta = .17$).

Significant F-tests were followed by a posteriori contrasts conducted using the Bonferroni-Holm procedure with corrected alpha levels (alphas = .016, .025 and .050) to reduce the likelihood of Type I error. Their results revealed that the Self-regulation strategies scores of students in G1 were significantly higher than those of G3 (no-question control group) (mean difference = .464; $p<.014$), but not significantly better than those of G2 (question-generation without any explicit instruction) (mean difference = .188, $p>.466$). Moreover, the Metacognitive knowledge scores of students in G1 were significantly higher than those of G2 (mean difference = .554, $p<.001$), but not better than those of G3 (mean difference = .188, $p>.466$). Effect size (i.e., practical significance) was, according to Cohen's (1988) criteria, medium to large (.14) for Self-regulated strategy use and large for Metacognitive knowledge (.17).

## Discussion

This study investigated the impact of teaching students to generate questions on how they learn and study in science.

Consistent with our prediction, results indicate that the intervention has an overall significant effect and that G1, the students trained to generate questions by providing prompts, showed the highest average gain, from pre-test to post-test, on Metacognitive knowledge and Strategic self-regulation. Our results appear to confirm that the effectiveness of self-questioning lies to a large extent in its metacognitive characteristics (Davey & McBride, 1986; King, 1992; Wong, 1985), which are critical for generative learning (Landa, 2000). Moreover, these results are somewhat in line with earlier research evidence linking students' question-generation training to improvements in their text processing (e.g., Bugg & Daniel, 2012; Rosenshine et al., 1996; Weinstein et al., 2010; Wong, 1985) and comprehension of lectures (King, 1992; 1994). However, our study is new in that these findings extend from the proximal domain of text processing and lectures to the distal domain of students' perceptions of how they learn and study, specifically to the two components of metacognition: (i) metacognitive knowledge (declarative, procedural and conditional) and (ii) regulation (i.e., planning, monitoring, and evaluation).

The results on the value of prompts are partially consistent with our original prediction. The ordered direction of between-groups differences in self-regulated strategy use was anticipated, but these are only statistically significant when comparing G1 and G3. Regarding Metacognitive knowledge, the direction is only partially as expected (G1 > G2 < G3) and differences are only statistically significant when comparing G1 and G2. These results seem to only partly support the claimed value of prompts in improving learning (e.g., Rosenshine et al., 1996), not surprising in view of Bugg and McDaniels's (2012, p. 922) warning: "experimental investigations, however, have not uniformly supported this claim, likely in part because […] a wide range of methods have been used to stimulate self-generation of questions […] the results in the literature vary […] across prompt type". Our findings add that the value of prompts seems also to depend on the type of DV being examined. Three possible factors might explain our results: (a) students in the control group had to clarify vocabulary. This activity may have induced them to enhance their awareness about the importance of clarifying for learning, which is one of the core elements of Reciprocal teaching (Palincsar & Brown, 1984); (b) the brevity of the intervention decreased the probability of detecting strong effects on learning; and (c) our DVs are distal rather than proximal to the domain of text processing and lectures, which does not increase that probability.

Taken together, these findings suggest that (a) the intervention on question-generation has an overall significant impact on how students learn and study; (b) this impact is focused on changes in metacognition (metacognitive knowledge and regulation) and has practical significance, in spite of the brevity of the intervention; (c) these changes are somewhat related to the influence of procedural prompts as generic question stems; and (d) teachers could help their students improve how they perceive their learning and study in science by training them in the use of these prompts.

Although these findings suggest a tendency, some limitations and areas for future improvement and research should be mentioned. First, the sample size was relatively small, the intervention was brief and the consequences of both circumstances probably decreased group differences and effect size estimates. Second, the activity completed by the control group might have favoured their results and somewhat unbalanced comparisons with the other groups. Future studies should choose for the control group an activity unrelated, as far as possible, to learning, engage a large sample of schools and increase the length of the intervention. These improvements would probably extend the impact of the intervention to learning approaches and question-asking in class. Finally, the use of on-line systems (see Núñez et al., 2011; Yu, 2009) opens new windows to design motivating and customisable scaffolded systems for teaching students to generate questions.

## Acknowledgements

This research was funded by the Spanish Ministry of Science and Innovation (MICINN, EDU2011-27416).

## References

Appleton, K., & Beasley, W. (1994). Students' learning in science lessons: Towards understanding the learning process. *Research in Science Education, 24,* 11-20.

Arias, A.V., Cabanach, R.G., Núñez, J.C., & González-Pienda, J.A. (1998). Cognitive-motivational variables, approaches to learning, and academic achievement. *Psicothema, 10,* 393-412.

Ato, M., & Vallejo, G. (2007). *Diseños experimentales en Psicología* [Experimental designs in Psychology]. Madrid: Pirámide.

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review, 5,* 243-260.

Bugg, J.M., & McDanield, M.A. (2012). Selective benefits of question self-generation and answering for remembering expository text. *Journal of Educational Psychology,* 922-931.

Chin, C., & Brown, D.E. (2000). Learning in Science: Comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*, 109-138.

Chin, C., & Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in Science Education*, *44*, 1-39.

Chin, C., Brown, D.E., & Bruce, B.C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education, 24*, 521-549.

Craig, F., & Lockhart, R.S. (1972). Levels of processing: Framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

D'Agostino, R.B., & Pearson, E.S. (1973). Tests for departure from normality. Empirical results for the distributions of b2 and √b1. *Biometrika, 60*, 613-622.

Davey, B., & McBride, S. (1986). Effects of question-generation on reading comprehension. *Journal of Educational Psychology, 78*, 256-262.

Entwistle, N., & McCune, V. (2004). The conceptual bases of study strategy inventories. *Educational Psychology Review, 16*, 325-345.

Graesser, A.C., & Person, N.K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*, 104-137.

Jarque, C.M., & Bera, A.K. (1987). A test for normality of observations and regression residuals. *International Statistical Review* , *55*, 163-172.

Kember, D., Biggs, J., & Leung, D.Y.P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology, 74,* 261-280.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, *30,* 338-368.

King, A. (1992). Comparison of self-questioning, sumarizing and notetaking-review as strategies for learning from lectures. *American Educational Research Journal, 29*, 301-323.

Land, S.M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, *48*, 61-78.

Lee, H.W, Lim, K.Y., & Grabowski, B.L. (2007). Generative learning: Principles and implications for making meaning. In J.M. Spector, M.D. Merrill, J.V. Merrienboer, & M.P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 111-124).

Marton, F., & Säljo, R. (1997). Approaches to learning. In F. Marton, D. Hounsell, & N.J. Entwistle (Eds.), *The Experience of Learning. Implications for Teaching and Studying in Higher Education* (pp. 39-58). Edinburgh: Scottish Academic Press.

National Research Council (2012). *A Framework for K-12 Science Education: Practices,Crosscutting Concepts, and Core Ideas*. Washington DC: The National Academies Press.

Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw Hill.

Núñez, J.C., Cerezo, R., Bernardo, A., Rosário, P., Valle, A., Fernández, E., & Suárez, N. (2011). Implementation of training programs in self-regulated learning strategies in Moodle format: Results of a experience in higher education. *Psicothema*, *23*, 274-281.

Otero, J., & Graesser, A.C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction, 19,* 143-175.

Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *2*, 117-175.

Pedrosa de Jesus, H., Almedia, P.A., Teixeira-Dias, J., & Watts, M. (2006). Students' questions: Building a bridge between Kolb's learning styles and approaches to learning. *Education and Training, 48*, 97-111.

Pintrich, P.R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385-407.

Pressley, M., Borkowski, J.G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of Child Development* (Vol. 5, pp. 89-129). Greenwich, CT: JAI Press.

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.

Shell, D.F., Husman, J., Turner, J.E., Cliffel, D.M., Nath, I., & Sweany, N. (2005). The impact of computer-supported collaborative learning communities on high school students' knowledge building, strategic learning, and perceptions of the classroom. *Journal of Educational Computing Research*, *33*, 327-349.

Sperling, R., Howard, L., Miller, L., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27,* 51-79.

Weinstein, Y., McDermott, K.B., & Roediger, H.L. (2010). A comparison of study strategies for passages: Re-reading, answering questions, and generating questions. *Journal of Experimental Psychology: Applied, 16*, 308-316.

Wittrock, M.C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 354-376.

Wong, B.Y.L. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, *55*, 227-268.

Yu, F.Y. (2009). Scaffolding student-generated questions: Design and development of a customizable online learning system. *Computers in Human Behavior, 25*, 1129-1138.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 1-37). Mahwah, NJ: Erlbaum.

Zimmerman, B.J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*, 166-183.