edual editorial UNIVERSIDAD DE ALMERÍA

Electronic Journal of Research in Educational Psychology ejrep

UNIVERSIDAD DE ALMERÍA

Ilustre Colegio Oficial de Psicología DE ANDALUCÍA ORIENTAL

Does Computer-based Elaborated Feedback Influence the Students' Question-Answering Process?

Ignacio Máñez[1], Eduardo Vidal-Abarca[1], Tomás Martínez[1]

[1] Department of Developmental and Educational Psychology, ERI-Lectura, University of Valencia, Valencia

Spain

Correspondence: Ignacio Máñez. Departamento de Psicología Evolutiva y de la Educación. Facultad de Psicología. Avda. Blasco Ibañez, 21. 46010 Valencia. Spain. E-mail: Ignacio.Manez@uv.es

© Universidad de Almería and Ilustre Colegio Oficial de la Psicología de Andalucía Oriental (Spain))

Abstract

Introduction. Students usually answer comprehension questions from texts as part of their academic activities. Elaborated Feedback (EF) has been found to be relatively effective to improve comprehension and learning from texts. However, there is little research on how computer-based feedback influences the question-answering process. This study examines the effects of a specific EF on the students' question-answering performance and their accuracy to assess textual relevance. Further, the study explores whether this EF influences the question, the text-search, and the feedback processing.

Method. Seventy-five 7th and 8th grade students answered a set of 20 questions from two texts (10 questions per text). During the question-answering process, students had the text available and were forced to highlight the text information they considered relevant to answer each question. While half the students received item-based EF that included information on the student's answer correctness and her accuracy to select question-relevant text information along with monitoring hints on task-specific strategies, the other half received non-formative feedback (i.e., control feedback group).

Results. EF enhanced the students' question-answering performance and reduced the amount of non-relevant text information students assessed as question-relevant. However, EF did not affect the question-relevant text information highlighted. Although EF did not influence the question and the text-search processing times in comparison to the control feedback group, students who received EF devoted more time and accessed additional information more often than the control group.

Discussion and Conclusion. Computer-based EF on the students' question-answering performance and their accuracy to assess textual relevance can be efficient to improve the students' question-answering process. This study sheds light on how to deliver EF in a digital environment. However, further research is necessary to explore the cognitive and metacognitive processes involved in feedback processing.

Keywords: elaborated feedback (EF); question-answering tasks; textual relevance; formative feedback; digital learning environments.

Resumen

Introducción. Los estudiantes suelen responder preguntas de comprensión a partir de textos como parte de sus actividades académicas. La retroalimentación elaborada (EF) es relativamente eficaz para mejorar la comprensión y el aprendizaje a partir de textos. Sin embargo, existe poca investigación sobre cómo la retroalimentación influye en el proceso de responder preguntas en entornos digitales. Este estudio examina los efectos de una retroalimentación elaborada específica sobre el rendimiento de los estudiantes al responder preguntas y su precisión para evaluar la relevancia textual. Además, el estudio explora si esta EF influye en el procesamiento de la pregunta, la búsqueda en el texto y la retroalimentación.

Método. Setenta y cinco estudiantes de 1º y 2º de Educación Secundaria respondieron un conjunto de 20 preguntas de dos textos (10 preguntas por texto). Durante el proceso de responder las preguntas, los estudiantes tuvieron acceso al texto, ya que tuvieron que seleccionar la información textual que consideraban relevante para responder cada pregunta. Después de responder cada pregunta, la mitad de los estudiantes recibió EF que incluía información sobre la corrección de sus respuestas y su precisión para seleccionar información relevante junto con sugerencias para monitorizar estrategias específicas de la tarea. La otra mitad no recibió retroalimentación formativa (i.e., grupo de retroalimentación control).

Resultados. La EF mejoró el rendimiento de los estudiantes al responder las preguntas y redujo la cantidad de información textual no relevante que los estudiantes evaluaron como relevante. Sin embargo, la EF no afectó a la información textual relevante seleccionada para cada pregunta. Aunque la EF no influyó en los tiempos de procesamiento de la pregunta y la búsqueda de información textual en comparación con el grupo de retroalimentación control, los estudiantes que recibieron EF dedicaron más tiempo y accedieron a información adicional más frecuentemente que el grupo de control.

Discusión y **Conclusion**: La EF sobre el rendimiento de los estudiantes al responder las preguntas y la precisión al evaluar la relevancia textual puede ser eficiente para mejorar el proceso de responder preguntas. Este estudio informa sobre cómo proporcionar EF en un entorno digital. Sin embargo, es necesario realizar más investigaciones para explorar los procesos cognitivos y metacognitivos involucrados en el procesamiento de la retroalimentación.

Palabras clave: retroalimentación elaborada (EF); tareas de responder preguntas; relevancia textual; retroalimentación formativa; entornos digitales de aprendizaje

Introduction

In school settings, one of the most frequent activities aimed at learning conceptual information refers to answering comprehension questions from an available text (Ness, 2011). Likewise, international assessments of reading skills like PISA (e.g., OECD, 2010) follow an open-text assessment in which students are expected to use written documents to perform a question-answering task. These scenarios require students not only to comprehend textual information (i.e., construct meaning), but also to assess textual relevance to answer the questions. Finding effective instructional procedures to enhance these reading skills is of most interest for the educational community. Although teachers usually know the students' responses to the questions, they may struggle to know whether each and every student has searched the text and how accurately they have assessed the textual relevance. Thus, teachers may deliver corrective feedback on the students' responses, but they rarely deliver more elaborative feedback. According to Shute (2008), corrective feedback includes Knowledge of Response (KR) and Knowledge of Correct Response (KCR) feedback. Whereas KR feedback is made of verification information (e.g., *Correct* or *Incorrect*), KCR feedback also provides information about what the correct answer was (e.g., *the correct choice depicted in green and the wrong choices depicted in red*). Elaborated Feedback (EF), on the other hand, may take many different forms since this feedback embraces a wide variety of messages (e.g., hints, explanations or examples), which are usually provided along with KR and KCR feedback.

Digital learning environments are able to overcome teacher's limitation regarding the lack of knowledge on how the students search the text for relevant information. Computer-based systems can store and grade meaningful variables automatically while students perform the task. Imagine a student answering a comprehension question from an available text in a computer. When the student accesses the text, she is expected to search for question-relevant text information, which means that she has to discriminate between more and less relevant information. If the computer-based system forces the student to indicate what information is germane to the question, it is possible to provide her with EF not only on her question-answering performance but also on her accuracy to assess textual relevance. In the present study, item-based EF included information on the student's answer (KR and KCR feedback) and her accuracy to select question-relevant text information (KR feedback and optionally-delivered KCR feedback), as well as specific monitoring hints on task-specific strategies (e.g., "*Review the text to see what information you missed. Next time try to select all the relevant information*"). In order to facilitate feedback processing, EF was delivered in three different

screens (see the Method section for a detailed description). This EF is likely to improve the students' comprehension skills. Prior research has analyzed the effects of feedback on text comprehension and the transfer of reading skills (e.g., Golke, Dörfler, & Artelt, 2015; Llorens, Vidal-Abarca, & Cerdán, 2016). However, no study has explored the feedback effects on the question-answering process (e.g., students' accuracy to assess textual relevance, and the time devoted to process the question and the text content). Likewise, there is little research on the use of feedback information in digital environments. In the present study, we aimed at answering the research questions "Can this EF improve students' question-answering performance and their accuracy to assess textual relevance?" and "Does this EF influence students' attention to process the question, the text and the feedback over a control feedback?" In what follows, we analyze the processes involved in answering comprehension questions from an available text and the role of feedback in question-answering tasks.

Question-answering tasks

In question-answering scenarios, students are expected not only to understand the text content, but also to use the text information strategically to solve the tasks (Gil, Martínez, & Vidal-Abarca, 2015; Vidal-Abarca, Mañá, & Gil, 2010). Students have to construct meaning by means of deploying both basic (e.g., word decoding) and higher-order (e.g., bridging or prior knowledge inferences) comprehension processes. Besides, the students have to monitor both their understanding and the textual relevance, which means that they consciously supervise their own comprehension and how relevant some pieces of text information are. Thus, the students monitor the text processing by making decisions (Vidal-Abarca et al., 2010) related to *when* and *what* text information they have to search as they answer the questions (Llorens et al., 2016). Regarding the text-search process, Vidal-Abarca et al. (2010, experiment 2) examined the students' decision to search or not to search the text while answering a set of comprehension questions. For each question, Secondary Education students had to assess how sure they were to provide the correct answer without searching the text. Main findings suggested that students tended to make no-search decisions when they were sure or quite sure to provide the correct answer. Results also showed that text-search monitoring accuracy had an impact on question-answering performance. When searching the text, evidence from eye-movement studies suggests that readers focus their visual attention on different parts of the text since there are text elements more and less relevant for each particular task (Kaakinen & Hyönä, 2011).

Question-answering tasks can be envisaged as instructions that direct students' attention to the question-relevant text information. The impact that task instructions have on processing task-relevant and non-relevant text information is known as the *relevance effect* (McCrudden & Schraw, 2007). Once the student has read the question, she has to assess the need to engage in searching the text to answer the question. If she decides to answer the questions without (re)processing at least some pieces of text information, she would have to rely on her initially-built representation to give an answer. However, if she engages in processing the text, she would have to locate the pieces of information that are relevant to the question. Thus, the student initiates a number of self-regulated behaviors aimed at providing the correct answer, so that the student has to distinguish between relevant and less-relevant or irrelevant information to the question (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; McCrudden & Schraw, 2007). Once the student is satisfied or her answer cannot be improved, she would stop processing the text and give an answer. These strategic text-search behaviors can be traced by computer-based systems (Cerdán, Vidal-Abarca, Martínez, Gilabert, & Gil, 2009; Vidal-Abarca et al., 2010). Previous evidence on the impact of searching the text suggests that both the number of search decisions and accurate no-search decisions play a positive and significant role in explaining performance scores in question-answering tasks (Gil et al., 2015). These strategic behaviors can be used for assessing (and maybe teaching) students' reading literacy skills (Rouet, 2006).

It is well-known that some students have difficulties to build a coherent text-level representation (Coté, Goldman, & Saul, 1998; Graesser, McNamara, & Louwerse, 2003). Among other variables, students' cognitive and metacognitive skills influence the students' ability to construct meaning from texts. Whereas some students may struggle to achieve comprehension because of poor cognitive skills (e.g., making inferences), others may struggle because of poor metacognitive skills (e.g., monitoring her comprehension) (e.g., Cain, Oakhill, Barnes, & Bryant, 2001). For instance, prior evidence suggests that less-skilled comprehenders experience difficulties to generate inferences (Cain & Oakhill, 2007), and to monitor comprehension accurately (Oakhill, Hartt, & Samols, 2005).

Further, students' ability to assess textual relevance is clearly influenced by comprehension skills since it depends on how accurate the students understand both the text and the question content. Overall, skilled comprehenders are more flexible with their reading behaviors (Martínez, Vidal-Abarca, Gil, & Gilabert, 2009) and are more strategic searchers than less-skilled comprehenders when they are asked to find question-relevant information (Catal-

do & Oakhill, 2000). For instance, Vidal-Abarca et al. (2010, experiment 1) and Gil et al. (2015) found that skilled comprehenders gave the correct answer just after processing a piece of question-relevant text information more often than less-skilled comprehenders. Thus, the authors concluded that skilled comprehenders monitored the use of relevant information more accurately than less-skilled comprehenders. Cerdán, Gilabert, and Vidal-Abarca (2011) showed how both skilled and less-skilled comprehenders assess textual relevance initially misguided by surface cues like word matching between the question and the text content. However, the skilled comprehenders are able to overcome that limitation and look for question-relevant information based on deeper semantic cues. Ramos and Vidal-Abarca (2013) had skilled and less-skilled comprehenders think-aloud while answering a set of questions from a text. Findings showed that skilled comprehenders reported more accurate verbalizations regarding textual relevance assessments than less-skilled comprehenders. Since we aimed at examining how accurate students assess textual relevance, as well as how EF may influence those assessments, we forced them to search and select the text information they considered relevant to answer each question.

The role of formative feedback in question-answering tasks

Feedback has been used in many different learning scenarios under a wide range of conditions with the intention of improving student's understanding or skills on the task performed. However, research on feedback effectiveness has shown inconsistent findings (e.g., Kluger & DeNisi, 1996; Shute, 2008), especially when it comes to higher-order processing tasks. To be effective, feedback content must be precise and interpretable by the students, guiding them in monitoring their strategies (Hattie & Timperley, 2007). While teacher-delivered feedback has some limitations (e.g., includes general information based on outcomes or it is not timely delivered), computer-based systems are capable of delivering adaptive feedback automatically based on a number of parameters related to the student's execution. Thus, students may compare their performance with a standard, which seems more effective than just providing the correct answer (e.g., Hattie & Timperley, 2007; Mason & Bruning, 2001). Regarding complexity, Shute (2008) distinguished between Knowledge of Results or KR (i.e., verification information about the correctness of a response: *correct* or *incorrect*), Knowledge of Correct Response or KCR (i.e., information about what the correct answer was), or Elaborated Feedback or EF (i.e., any additional information beyond the student's performance such as explanations, prompts or examples). The latter type of feedback usually includes simple KR or KCR feedback. Recently, van der Kleij, Feskens, and Eggen (2015)

conducted a meta-analysis on the feedback effectiveness within computer-based formative assessments in which EF was found to be the most effective type of feedback over KCR and KR (mean effect sizes were .49, .32, and .05, respectively). This positive effect of EF may be explained by the probability of reducing the gap between the inaccurate or erroneous performance and the standard performance.

Question-answering tasks can include feedback information to improve students' comprehension and their accuracy to assess textual relevance as long as the feedback messages include precise information on the task performed and students are willing to invest some effort in processing them. To the best of our knowledge, only a few studies have explored whether computer-based formative feedback influences student's text comprehension (e.g., Golke et al., 2015; Lee, Lim, & Grabowski, 2009) and reading strategies (e.g., Llorens et al., 2016; Llorens, Vidal-Abarca, Cerdán, & Ávila, 2015). For instance, Golke et al. (2015) examined the effects of different types of EF (i.e., KR feedback along with inference-prompts, error explanations, and monitoring-prompts) in comparison to KR feedback and no feedback on sixth-grade students' text comprehension performance. Feedback was delivered after providing incorrect responses. Participants had a second attempt for those incorrectly answered questions, although they did not receive feedback on this attempt. Participants took a post-test and a follow-up test. Findings showed that neither the student's comprehension performance nor the time devoted to answer the questions differed across the types of feedback. In other words, none of the types of computer-based EF increased performance or influenced the time on task in comparison to the KR or the no feedback control groups. The authors suggested that students may not have engaged actively in processing computer-based feedback to enhance text comprehension since they did not spend much time to fix initial errors. Golke et al. (2015) did not assess the students' accuracy to distinguish between relevant and non-relevant text information as an effect of EF, which is one of the goals of the present experiment.

Previous evidence suggests that providing students with EF that indicates where the relevant information is located increases the students' text-search decisions over an EF made of prompts to search the text (Llorens et al., 2015). In this study, the computer-based system traced the students' text-search behaviors and delivered EF based on these actions. Further, Llorens et al. (2016, experiment 1) studied the effects of EF on the transfer of self-regulated reading strategies to new question-answering tasks in which no feedback was delivered. EF included KR, KCR and tailored hints on how to proceed in the next questions based on students' current performance. It should be noted that whereas one group of students were al-

lowed to search the text freely during the question-answering process, another group of students had to search the text and select question-relevant text information. Results suggested that selecting question-relevant information and receiving EF increased the students' use of strategic reading skills (i.e., when to search and what information to search) and their comprehension performance on a post-test. Additionally, Llorens et al. (2016, experiment 2) found that both selecting question-relevant text information and receiving EF played independent roles in the students' text comprehension and the transfer of strategic reading skills. However, the authors did not examine either the effects of EF on the students' accuracy to assess textual relevance or its influence on the time devoted to process the question and the text, which is one of our research goals in this paper.

To be effective, students must engage actively in processing the feedback (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). However, there is a lack of research on the effects of feedback on the students' question-answering process (i.e., time processing the question, searching the text, and processing the feedback) (Golke et al., 2015). One of the main advantages of computers is that students are given the possibility of making ongoing decisions to use the materials, so that some pieces of feedback can be delivered upon students' request. In a pioneering study, Timmers and Veldkamp (2011) had university students answer a set of questions on information literacy and received an overview of KR feedback for each answer. Afterwards, they had the option to voluntarily access EF made of KCR and an explanation. Results indicated that the students' decision to access the EF varied greatly among students. In our case, students were allowed to access specific feedback on the standard question-relevant text information. Thus, we are able to examine the extent to which students are willing to engage in processing feedback.

Objectives and Hypotheses

The present study has two main goals:

1. To determine the impact of a specific EF on students' question-answering performance and their accuracy to assess textual relevance. We predict that EF would improve students' question-answering performance and their accuracy to assess textual relevance (i.e., students would select more relevant and less non-relevant text information) in comparison to a control feedback group.

2. To examine the extent to which EF influences the question, the text-search, and the feedback processing. We predict that EF would not affect the time students devote to process either the question content or the text information (Golke et al., 2015). However, EF would increase the processing time and the students' decisions to access optionally-delivered feedback over a control condition.

Method

Participants

A total of seventy-five 7th and 8th grade students aged 11-15 (*M* = 12.33; *SD* = .704) completed the whole study without incidences (46.7% were female). All participants were Spanish native speakers. The data of seven students were excluded from the data analysis due to absences during one or more phases of the study (n = 3) or software failures (n = 4). Participants were randomly assigned to one out of two groups according to the type of feedback (feedback type: EF, Control) after measuring students' prior comprehension skill as their ability to answer comprehension questions from an available text. Participants' confidentiality was preserved and the School's boards approved the application of the study.

Instruments

Test of Reading Literacy for Secondary Education (CompLEC; Llorens et al., 2011). We initially used CompLEC, a standardized test to assess participants' prior comprehension skill. This is a paper-and-pencil test that includes five texts and twenty comprehension questions developed according to the PISA framework (OECD, 2006). Each correct response received a score of 1 and each incorrect a zero (maximum score: 20). The psychometric properties of CompLEC revealed acceptable reliability (*Cronbach's alpha* = .80) (Llorens et al., 2011).

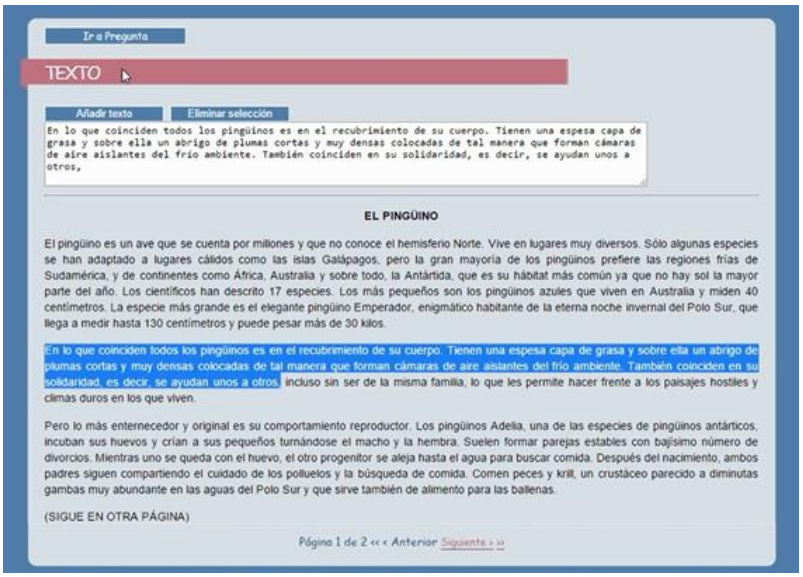
Test of Comprehension Processes (TCP; Martínez, Vidal-Abarca, Sellés, & Gilabert, 2008). For the testing phase, we employed the TCP, a standardized test made of two expository texts (i.e., Penguins and Sioux with a length of 550 and 471 words, respectively) and 10 multiple-choice comprehension questions per text. The topic of the texts is not covered in the academic curriculum, so that students have little to no prior knowledge. Regarding readability indexes, the Flesch-Szigriszt reading grade levels were 66.89 for Penguins and 63.39 for Sioux, indicating that these texts can be considered "Quite easy" and "Normal" to read ac-

cording to the INFLESZ Scale. The semantic relations between text sentences were analyzed with ETAT (Expository Text Analysis Tool; Vidal-Abarca, Gilabert, & Abad, 2002; Vidal-Abarca, Reyes, Gilabert, Calpe, Soria, & Graesser, 2002). These analyses showed which fragments of the text information should be considered relevant to answer each question. The relationship between each question and the relevant text information has been empirically validated using a large sample of students of similar age to the ones in the present study (Martínez et al., 2009). Text segments with question-relevant information could be located together or separated by other sentences. Each question had four answer choices, with only one correct response, so that correct responses received a score of 1 and incorrect responses a zero (maximum score: 20). The psychometric properties of the computer-based TCP revealed moderate indices of reliability (*Cronbach's alpha* = .72) and validity (Pearson correlation between total scores obtained on the computerized version of the TCP and scores on the comprehension subtest of PROLEC-SE = .64) (Martínez et al., 2009). According to the authors, these indices were slightly lower than the reliability (.79) and validity (.72) obtained for the paper-and-pencil TCP version due to the variance introduced by the "moving window" technique. Further, the TCP and the eCompLEC scores were found to be significantly correlated ($r = .53$, $p < .01$) (Gil et al., 2015).

Read&Learn. Participants completed both tasks (i.e., answering 10 questions per text and selecting the question-relevant text information) on Read&Learn, a computer-based system based on Read&Answer technology (Vidal-Abarca et al., 2011). Read&Learn is able to store online behaviors and processing times throughout the task completion in order to deliver feedback based on the student's execution and performance. Texts and questions were presented on different screens (Figures 1a and 1b). Each text was presented in two different pages because of its length. Before accessing the first question, students were presented with the text information using the masking/unmasking technique for the initial text reading, so that they had to click on each piece of text information that they wanted to read in order to make its content visible. Only one piece of text information was visible to read at a time. Same masking procedure was used for the multiple-choice questions. Once the students decided to unmask the question wording, the choices remained masked, and vice versa (Figure 1b). However, due to the nature of the selection task, the text information remained fully visible once students answered the first question. Students were instructed to select the relevant text information for each question by clicking and dragging the mouse to select the text infor-

mation they considered germane to the question and clicking the button “Add text”. Students were allowed to change and adjust their selection by clicking the button “Delete selection” and repeating the process (Figure 1a). Read&Learn system allowed them to select words, sentences, or entire paragraphs. Please note that the information selected on the text screen was displayed automatically on the question screen. Participants were allowed to move back and forth between the text and the question screens by means of the “Text” and “Question” buttons located in the upper part of the question and the text screens respectively. This interface feature provided flexibility to the question-answering process.

a) Text screen



b) Question screen

Figure 1. Screenshots of the task: (a) the text screen with a piece of text information selected, and (b) the question screen, with the question stem unmasked (above) and the answer choices unmasked (below).

After selecting the question-relevant text information and choosing a response, the system delivered a set of feedback messages depending on the condition assigned. For the EF group, participants initially received a written message in a pop-up box. This message was made up of KR on the correctness of the student's answer ("*Correct*" or "*Incorrect*") and their accuracy to select question-relevant information (e.g., "*You have selected only part of the necessary information*"), as well as specific monitoring hints on task-specific strategies (e.g., "*Review the text to see what information you missed. Next time try to select all the relevant information*") (Figure 2a).

Once they closed that message, the question content with the students' selection was displayed on screen with visual feedback information. They received KCR on the answer to the question and KR on their selection, showing the correct choice in green and the wrong choices in red, as well as the relevant pieces of text information depicted in green and the non-relevant pieces of text information in red (Figure 2b). Then, participants had the option to access the text screen with the standard question-relevant text information highlighted, so that KCR feedback on the standard relevant information was delivered (Figure 2c).

For the Control feedback group, students performed the same tasks but received non-formative feedback messages. Participants initially received a written message in a pop-up box (e.g., "*You have answered question number 1*"). Once they closed that message, the question content with the students' selection was displayed on screen without any visual feedback

information. Then, participants had the option to access the text screen, although no feedback on the standard question-relevant text information was delivered. Whereas the first two screens were delivered automatically, the third one was delivered upon students' request. All the participants were allowed to move to the next question by clicking on the "Next Question" button displayed on screens b and c.

a)

b)

c)

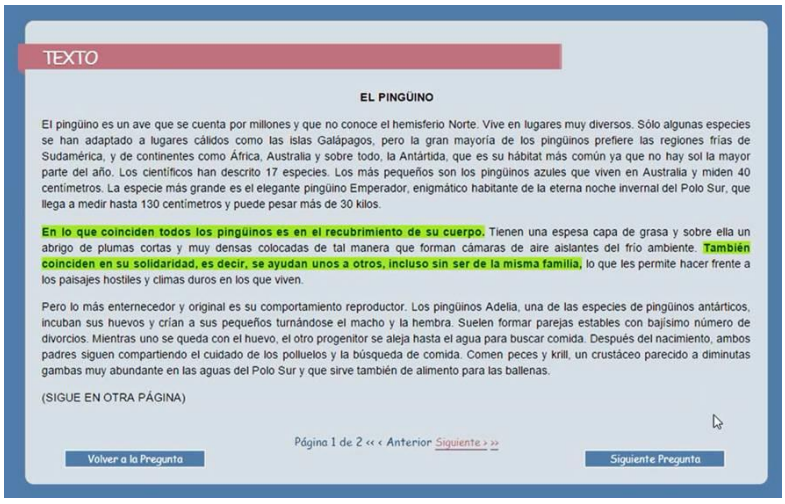


Figure 2. Screenshots of the EF messages: (a) pop-up box with written feedback, (b) KCR on the answer to the question and KR on the students' selection, and (c) text review screen with KCR on the standard question-relevant text information.

Procedure

Students were assessed in two sessions of approximately 1 hour each. In the first session, participants' prior comprehension skill was assessed with the standardized test CompLEC (Llorens et al., 2011). CompLEC was administered in their regular classroom to make sure that both groups were similar in comprehension skill. In the second session, students completed the tasks in the computer-lab classroom. They were first instructed on how to use Read&Learn's interface to read the texts, select the question-relevant text information and answer the questions, as well as how to proceed when computer-based feedback was delivered. Students were informed to read the text before accessing the first question in order to reduce the variability of initial text reading among students. Once students had read the text, they moved to the first question. They were allowed to move back and forth between the text and the question screens since they had to select the question-relevant text information to answer the question. After performing both tasks, participants received feedback. Questions were presented one at a time, so that this task procedure was repeated for each question. Participants performed the task individually in the computer room while their behaviors were stored and graded automatically by the computer-based system Read&Learn.

Data Analysis

We conducted a number of analyses of covariance (ANCOVA) using SPSS v.24 for each dependent variable. The study was designed according to the quantitative methodology. Read&Learn computed and stored a number of quantitative measures for both the question-answering and the selection tasks: (a) *Question-answering performance* was the proportion of correct responses (e.g., 50% indicates that the student provided the correct answer in 10 out of the 20 questions), (b) *Relevant information* was the proportion of relevant information selected over the standard question-relevant text information, (c) *Non-relevant information* was the proportion of irrelevant information selected over the student's complete selection. To illustrate how Read&Learn computed these measures we examine the example provided in Figures 1 and 2. This student selected almost all the relevant text information. Read&Learn com-

puted the percentage of relevant information selected by means of comparing the words included in the student's selection (i.e., Figure 1) against the standard selection (i.e., information provided in Figure 2c). In this case, the student selected about 80% of the standard selection, which indicates she was pretty accurate to find the question-relevant information. However, this student also selected non-relevant information. Read&Learn computed the percentage of non-relevant information selected by means of comparing non-relevant words included in the student's selection against the student's complete selection (i.e., Figure 2b). In this case, the student's selection included about 50% of non-relevant information. These measures allowed us to distinguish in a precise way what pieces of the student's selection met the standard criteria and what pieces of her selection did not meet the standard. Regarding online processing measures, Read&Learn recorded several quantitative measures as well: (d) *Question reading time* was the amount of time students devoted to process the question stem and the answer choices, (e) *Text-search processing time* was the amount of time students spent searching and assessing textual relevance, (f) *Feedback reading time* was the amount of time students devoted to process the automatically-delivered feedback messages (i.e., first two feedback screens), (g) *Text review access* was the proportion of decisions to access the optionally-delivered text review screen, and (h) *Text review reading time* was the amount of time students spent processing the optional text screen.

Results

Effects of EF on the question-answering performance and the accuracy to assess textual relevance

To test our hypothesis 1, we computed three separate univariate ANCOVAs with the type of feedback (EF, Control) as independent variable, the students' prior comprehension skill as covariate, and *Question-answering performance* and accuracy to assess textual relevance (i.e., *Question-relevant* and *Non-relevant text information* selected) as dependent variables. The means and standard deviations are provided in Table 1.

As predicted, we found significant differences for the *Question-answering performance* as a function of type of feedback, $F(1, 72) = 5.11$, $p = .027$, *partial* $\eta^2 = .07$, indicating that students who received EF outperformed students who received Control feedback. For the accuracy to assess textual relevance, interesting findings were found. Contrary to our predictions, we found no significant differences for the *Question-relevant text information* selected as a function of type of feedback, $F(1, 72) = 0.24$, $p = .63$. For the *Non-relevant text infor-*

mation selected, we found significant differences as a function of type of feedback, $F(1, 72) = 10.09$, $p = .002$, *partial* $\eta^2 = .12$, indicating that students who received EF selected less non-relevant text information than students who received Control feedback. This set of results partially confirmed hypothesis 1. The covariate, students' prior comprehension skill, was significantly related to the *Question-answering performance*, $F(1, 72) = 34.32$, $p < .001$, *partial* $\eta^2 = .32$, the *Question-relevant information* selected, $F(1, 72) = 31.14$, $p < .001$, *partial* $\eta^2 = .30$, and the *Non-relevant text information* selected, $F(1, 72) = 27.28$, $p < .001$, *partial* $\eta^2 = .28$.

Table 1. *Means (Standard Deviations) of the Question-Answering Performance and the Textual Relevance Accuracy as a Function of Type of Feedback (EF, Control).*

	Q-A performance	*Relevant Text Information*	*Non-Relevant Text Information*
EF	.67 (.16)	.51 (.13)	.39 (.12)
Control	.59 (.21)	.52 (.17)	.47 (.13)

Effects of EF on the question, the text, and the feedback processing

To test our hypothesis 2, we computed a number of separate univariate ANCOVAs with the type of feedback (EF, Control) as independent variable, students' prior comprehension skill as covariate, and online processing measures (i.e., *Question reading time*, *Text-search processing time*, *Feedback reading time*, *Text review access* and *Text review reading time*) as dependent variables. The processing times were introduced in seconds. The means and standard deviations for the question reading time, the text-search processing time, and the feedback processing measures are provided in Table 2.

As predicted in hypothesis 2, no significant differences depending on the type of feedback were found neither for the *Question reading time*, $F(1, 72) = 2.59$, $p = .112$, nor for the *Text-search processing time*, $F(1, 72) = 0.37$, $p = .55$. Thus, both EF and Control feedback groups devoted the same amount of time to process the question and the text while searching for question-relevant information. The covariate, students' prior comprehension skill, was significantly related to the *Question reading time*, $F(1, 72) = 4.07$, $p = .047$, *partial* $\eta^2 = .05$, and the *Text-search processing time*, $F(1, 72) = 7.39$, $p = .008$, *partial* $\eta^2 = .09$.

Regarding feedback processing, we found significant differences depending on the type of feedback for the *Feedback reading time*, $F(1, 72) = 41.92$, $p < .001$, *partial* $\eta^2 = .37$, the number of *Text review accesses*, $F(1, 72) = 99.73$, $p < .001$, *partial* $\eta^2 = .58$, and the *Text review reading time*, $F(1, 72) = 15.52$, $p < .001$, *partial* $\eta^2 = .18$. These results showed that students who received EF spent more time processing those messages and accessed the optional text review screen more often than students who received Control feedback. These results are also consistent with hypothesis 2. For the feedback processing measures, however, students' prior comprehension skill was not significantly related neither for the *Feedback reading time*, $F(1, 72) = 0.47$, $p = .50$, nor for the *Text review accesses*, $F(1, 72) = 2.29$; $p = .13$, nor for the *Text review reading time*, $F(1, 72) = 0.43$, $p = .52$.

Table 2. *Means (Standard Deviations) of the Question Reading Time, the Text-search Processing Time, and the Feedback Processing Measures as a Function of Type of Feedback (EF, Control).*

	Question Reading Time	*Text-search Processing Time*	*Feedback Reading Time*	*Text Review Access*	*Text Review Reading Time*
EF	27.63 (6.80)	56.73 (16.37)	8.65 (2.51)	.43 (.22)	7.75 (4.06)
Control	30.10 (6.87)	54.49 (16.20)	5.44 (1.67)	.04 (.07)	3.20 (5.79)

Note. Processing times are measured in seconds; Text Review Access is measured in proportions

Discussion and conclusion

Digital learning environments are growing in popularity in school settings. Therefore, it is essential to know the extent to which training components such as formative feedback can influence students' performance and reading processes. This study builds on previous research on the effects of EF on text comprehension (e.g., Golke et al., 2015; Llorens et al., 2016) and the students' use of computer-based EF (e.g., Timmers & Veldkamp, 2011). Thus, the study reported in this paper investigates the effects of a specific EF on Secondary school students' question-answering performance and their accuracy to assess textual relevance. Further, this study examines the extent to which this EF influences the question, the text-search and the feedback processing over a control condition.

Main results indicated that (a) providing students with EF on both their question-answering performance and their accuracy to select question-relevant text information improved the students' comprehension scores and had a positive effect on avoiding non-relevant information, although it did not affect the question-relevant text information, (b) students who received the EF spent the same amount of time to process the question and the text information, and (c) students invested some extra time to process the EF information and decided to use the optionally-delivered feedback on the standard question-relevant text information.

As predicted in hypothesis 1, findings suggest that this specific tailored EF can improve students' text comprehension performance in two ways. First, EF enhanced students' question-answering performance over a control feedback that served as a baseline. This result is in line with previous findings on feedback effectiveness within question-answering tasks (e.g., Llorens et al., 2015; Llorens et al., 2016). Second, EF influences the amount of non-relevant text information students assess as question-relevant. Thus, students who were informed on their accuracy to assess textual relevance were able to reduce the non-relevant text information selected as relevant. Both findings indicate that computer-based EF can play a role on text comprehension because the messages provide valuable information to overcome comprehension problems (i.e., inaccurate or incorrect knowledge). Contrary to our first hypothesis, EF did not influence the students' accuracy to select question-relevant text information. Overall, these findings suggest that avoiding non-relevant text information may be a first step towards a more accurate assessment of question-relevant text information since selecting additional pieces of relevant information may involve complex processes that require extensive and explicit training. These findings confirm the importance of monitoring accurately the text-search process (i.e., what text information is relevant) when answering comprehension questions from an available text (Anmarkrud et al., 2013; McCrudden & Schraw, 2007; Vidal-Abarca et al., 2010) and how EF can exert an impact on it.

According to hypothesis 2 and in line with previous research (Golke et al., 2015), results showed that EF did not affect neither the time students devoted to process the question content nor the text-search processing time. Related to the students' accuracy to assess textual relevance, students who received EF were more efficient during the text-search process. In other words, they got better performance scores than the students in the control group for the question-answering task and the accuracy to select non-relevant text information while both

groups devoted the same amount of time to perform the task. As predicted in our hypothesis 2 on feedback processing, EF increased significantly the processing times and the students' decision to access the optionally-delivered feedback over our baseline made of non-formative information. In contrast to Golke et al.'s (2015) findings, it seems that students engaged in processing EF to compare their performance with the standard provided rather than just processing the correct response, which is essential to benefit students' learning (e.g., Hattie & Timperley, 2007; Mason & Bruning, 2001). Results showed that students decided to access more than 40% of the text review screens with the feedback on the standard question-relevant text information. Therefore, giving students the opportunity to access formative feedback voluntarily seems essential for the effect of feedback in digital learning environments (e.g., Llorens et al., 2016; Timmers & Veldkamp, 2011).

Further, it is interesting to acknowledge the role working memory and inhibition processes may have played on both students' text comprehension and feedback use since these executive functions are related to the readers' ability to comprehend textual information (Borella, Carretti, & Pelegrina, 2010; Cain, 2006), even when it comes to Secondary school students (Demagistri, Richards, & Juric, 2014). Working memory enables temporarily storing of information, as well as the integration of previously-activated knowledge from long-term memory with the newly-codified information from the text (Kendeou, Papadopoulos, & Spanoudis, 2012). Therefore, students with high working memory span may outperform their low span counterparts in text comprehension and even integrate feedback information with their knowledge in a deeper way. On the other hand, inhibition refers to the attentional control processes that actively suppress irrelevant task information that competes with relevant information within the working memory system. Thus, students with high inhibition ability may discard irrelevant task information accurately, not only when processing text but also feedback information. Findings of the present study suggest that delivering EF on the student's accuracy to select question-relevant information may have influenced their ability to inhibit irrelevant information to answer the questions. However, more research is necessary to examine the relationships among these executive functions, the reading comprehension processes and the feedback use in digital environments.

Besides that, this study has several limitations we plan to address in future studies. First, the students who received EF devoted a few more seconds in comparison to the control condition, so it is necessary to explore what cognitive and metacognitive processes EF trig-

gers when it is delivered, as well as what components of EF are mainly attended. To that end, the think-aloud methodology may provide valuable information. Second, current improvements in students' text comprehension and textual relevance assessments are dependent on immediate feedback. Future research should examine the transfer and long-term effects of formative feedback on the students' accuracy to assess textual relevance in new question-answering tasks. Finally, although we focused on one of the most common reading comprehension tasks used in school settings, it would be interesting to study the effects of EF on other learning domains such as science or mathematics.

Open questions aside, our results add to the research on the effects and use of computer-based EF when performing question-answering tasks and provide insights into how computer-based systems should deliver formative feedback aimed at enhancing students' learning. First, this specific EF seems to play a role in students' comprehension and ability to assess textual relevance. Thus, detailed EF messages have to be developed to promote students' improvements when performing high-order processing tasks. One of the main reasons of this positive effect may lie in the metacognitive prompt included in the messages to review the question and the text content. This element may have stimulated the students' decision to access the KCR feedback on the standard question-relevant text information, which in turn may have favored the students' accuracy to assess textual relevance. Second, findings suggest that students monitor and self-regulate the use of feedback, so that computer-based systems should give students the opportunity to access at least some information upon request. Thus, these results have relevant implications for any digital learning environment (e.g., e-textbooks, Intelligent Tutoring Systems like TuinLEC, or e-learning platforms like Moodle or MOOCs) that can deliver formative feedback as part of their training programs. Any of these digital environments are able to deliver assigned tasks and collect a bunch of useful data that can be transformed into tailored feedback.

References

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science, 41*(5), 873-894. doi:10.1007/s11251-013-9263-8

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213–238. doi:10.3102/00346543061002213

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities 43*(6), 541-552. doi:10.1177/0022219410371676

Cain, K. (2006). Individual differences in children´s memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory, 14*(5), 553-569. doi:10.1080/09658210600624481

Cain, K., & Oakhill, J. V. (2007). Reading comprehension difficulties: correlates, causes, and consequences. In K. Cain, & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41-75). New York, NY: Guilford.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*(6), 850-859. doi:10.3758/BF03196414

Cataldo, M. G., & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? An investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of Educational Psychology, 92*, 791-799.

Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011) Selecting information to answer questions: Strategic individual differences when searching texts. *Learning and Individual Differences, 21*, 201–205. doi:10.1016/j.lindif.2010.11.007

Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction, 19*(1), 13-27. doi: 10.1016/j.learninstruc.2007.12.003

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: relations between processing and representation. *Discourse Processes, 25*(1), 1-53. doi:10.1080/01638539809545019

Demagistri, M. S., Richards, M. M., & Juric, L. C. (2014). Incidencia del funcionamiento ejecutivo en el rendimiento en comprensión lectora en adolescentes. *Electronic Journal of Research in Educational Psychology, 12*(2), 343-370. doi:10.14204/ejrep.33.13146

Gil, L., Martínez, T., & Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education, 82*, 50-59. doi:10.1016/j.compedu.2014.10.026

Golke, S., Dörfler, T., & Artelt, C. (2015). The impact of elaborated feedback on text comprehension within a computer-based assessment. *Learning and Instruction, 39*, 123-136. doi:10.1016/j.learninstruc.2015.05.009

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet, & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). New York, NY: Guilford Publications.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. doi:10.3102/003465430298487

Kaakinen, J. K., & Hyönä, J. (2011). Online processing of and memory for perspective relevant and irrelevant text information. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 223–242). Greenwich, CT: Information Age Publishing.

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction, 22*(5), 354–367. doi:10.1016/j.learninstruc.2012.02.001

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. doi:10.1037/0033-2909.119.2.254

Lee, H. W., Lim, K. Y., & Grabowski, B. L. (2009). Generative learning strategies and metacognitive feedback to facilitate comprehension of complex science topics and self-regulation. *Journal of Educational Multimedia and Hypermedia, 18*(1), 5-25.

Llorens, A. C., Gil, L., Vidal-Abarca, E., Martínez, T., Mañá, A., & Gilabert, R., (2011). Evaluación de la competencia lectora: la prueba de Competencia Lectora para Educación Secundaria (CompLEC). *Psicothema, 23*(4), 808-817.

Llorens, A. C., Vidal-Abarca, E., & Cerdán, R. (2016). Formative feedback to transfer self-regulation of task-oriented reading strategies. *Journal of Computer Assisted Learning, 32*(4), 314-331. doi: 10.1111/jcal.12134

Llorens, A. C., Vidal-Abarca, E., Cerdán, R., & Ávila, V. (2015). Does formative feedback on search behavior help students in answering comprehension questions from an available text? *Infancia y Aprendizaje, 38*(4), 808-841. doi:10.1080/02103702.2015.1076269

Martínez, T., Vidal-Abarca, E., Gil, L., & Gilabert, R. (2009). On-line assessment of comprehension processes. *The Spanish Journal of Psychology, 12*, 308-319. doi:10.1017/S1138741600001700

Martínez, T., Vidal-Abarca, E., Sellés, P., & Gilabert, R. (2008). Evaluation of comprehension strategies and processes: Test of Comprehension Processes (TCP). *Infancia y Aprendizaje, 31*(3), 319-332.

Mason, B., & Bruning, R. (2001). Providing feedback in computer-based instruction: What the research tells us. Retrieved from https://www.researchgate.net/publication/247291218_Providing_Feedback_in_Computer-based_Instruction_What_the_Research_Tells_Us

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review, 19*, 113-139. doi:10.1007/s10648-006-9010-7

Ness, M. (2011). Explicit reading comprehension instruction in elementary classrooms: Teacher use of reading comprehension strategies. *Journal of Research in Childhood Education, 25*, 98–117. doi:10.1080/02568543.2010.531076

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing, 18*(7-9), 657-686. doi:10.1007/s11145-005-3355-z

OECD (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. París: OECD.

OECD. (2010). *PISA 2009 results: What students know and can do: Student performance in Reading, Mathematics and Science*. Paris: OECD Publishing.

Ramos, L., & Vidal-Abarca, E. (2013). Differences between students with high and low reading literacy skills: A study with think aloud methodology. *Cultura y Educación, 25*(3), 295-308. doi:10.1174/113564013807749722

Rouet, J. F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. doi:10.3102/0034654307313795

Timmers, C. F., & Veldkamp, B. P. (2011). Attention paid to feedback provided by a computer-based assessment for learning on information literacy. *Computers & Education, 56*, 923–930. doi:10.1016/j.compedu.2010.11.007

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475-511. doi:10.3102/0034654314564881

Vidal-Abarca, E., Gilabert, R., & Abad, N. (2002). A proposal for good expository text: Toward an expository text technology. *Infancia y Aprendizaje, 25*(4), 499-514. doi:10.1174/021037002762064064

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology, 102*(4), 817-826. doi:10.1037/a0020062

Vidal-Abarca, E., Martínez, T., Salmerón, L., Cerdán, R., Gilabert, R., Gil, L., Mañá, A., Llorens, A., & Ferris, R. (2011). Recording on-line processes in task-oriented reading with Read&Answer. *Behavior Research Methods, 43*(1), 179–192. doi:10.3758/s13428-010-0032-1

Vidal-Abarca, E., Reyes, H., Gilabert, R., Calpe, J., Soria, E., & Graesser, A. C. (2002). ETAT: Expository Text Analysis Tool. *Behavior Research Methods, Instruments, & Computers*, *34*(1), 93-107. doi:10.3758/BF03195428

Received: 28-09-2018
Accepted: 20-12-2018