



VNIVERSITATIS VALÈNCIA

Departamento de Psicología Evolutiva y de la Educación

**Procesamiento y efectividad de la retroalimentación
formativa para incrementar la comprensión y
aprendizaje de conocimiento conceptual en entornos
digitales**

Processing and effectiveness of formative feedback to increase
comprehension and learning of conceptual knowledge in digital
environments

TESIS DOCTORAL CON MENCIÓN INTERNACIONAL PRESENTADA POR:

IGNACIO MÁÑEZ SÁEZ

DIRECTORES:

EDUARDO VIDAL-ABARCA GÁMEZ

TOMAS MARTÍNEZ GIMÉNEZ

Programa de Doctorado: Lectura y Comprensión

Valencia, mayo de 2019

A mis padres.

AGRADECIMIENTOS

Aprovecho estas líneas para agradecer a todas las personas que de una forma u otra han contribuido a que esta tesis vea la luz tras varios años de trabajo.

A Eduardo y Tomás, mis directores, por ofrecerme la oportunidad de formar parte de vuestro equipo y por preocuparos por mi formación. A Eduardo, por todos los comentarios y consejos que han contribuido a la adquisición de conocimientos y competencias que aún no soy capaz de valorar adecuadamente, así como por tu apoyo en los momentos de incertidumbre que acompañan esta etapa de formación académica. A Tomás, por tus aportaciones científicas, visión crítica, y por ser el puente entre los miembros del grupo y los técnicos informáticos que han desarrollado Read&Learn, herramienta sin la que esta tesis no habría sido posible realizar. Gracias a ambos por contar conmigo, ofrecerme oportunidades, y confiar en mi cuando yo no era capaz de hacerlo. Y, como no, por vuestro ‘feedback’.

Gracias a todos los miembros del grupo Psicotext y de la ERI-Lectura con los que he tenido la oportunidad de formarme, por tenderme la mano y ayudarme a solucionar mis dudas. A Ramiro, por tu apoyo e innumerables consejos para gestionar la docencia. A Laura, Amelia y Vicen, con quienes he tenido la oportunidad de trabajar más directamente estos últimos años, por contar conmigo. A Raquel, Antonio, Lalo, Javi, Inma y Gemma, por vuestras aportaciones y consejos. A Carlos y Guillem, por crear Read&Learn y solucionar los problemas a tiempo. A los becarios de investigación con los que comencé mi andadura y a aquellos que se han ido sumando al grupo, con quienes he compartido dudas y buscado soluciones (Marian, Carmen, Pablo, Alba, Arantxa, y Nadina). Especialmente, quisiera agradecer a Marian, quien no ha dejado de darme buenos consejos desde que llegué al grupo.

También me gustaría agradecer a los investigadores que me abrieron las puertas de sus laboratorios para realizar estancias de investigación en el extranjero, especialmente a la Dra. Panayiota Kendeou (University of Minnesota, USA) y al Dr. Jean-François Rouet (Université de Poitiers, France), así como al resto de investigadores que de un modo u otro han aportado ideas a mis estudios durante los congresos y jornadas científicas en las que he tenido la oportunidad de participar.

Igualmente, me gustaría dar las gracias a todos los profesores, psicólogos y alumnos de los centros que han participado en la aplicación de los estudios aquí presentados.

Por supuesto, gracias a mi familia. A todas y cada una de las personas que han mostrado interés y se han preocupado por mí. Especialmente a mis padres, por vuestro cariño incondicional, por confiar en mí, por comprender lo incomprensible, y por apoyarme cuando más lo necesitaba.

Valencia, 11 de abril de 2019

Índice

INTRODUCTION.....	i
CAPÍTULO 1: MARCO TEÓRICO.....	1
1. EL PAPEL DE LAS PREGUNTAS PARA EVALUAR LA COMPRENSIÓN Y EL APRENDIZAJE.....	3
1.1. Marco de trabajo para las tareas de responder preguntas.....	5
1.2. El papel de las diferencias individuales al responder preguntas.....	17
2. RETROALIMENTACIÓN FORMATIVA PARA APRENDER CONOCIMIENTO CONCEPTUAL.....	22
2.1. Tipos y funciones de la retroalimentación formativa.....	25
2.2. Efectividad de la retroalimentación formativa.....	29
2.3. Marco de trabajo para el procesamiento de la retroalimentación.....	41
2.3.1. Supuestos del procesamiento de la retroalimentación.....	59
2.4. Diferencias individuales en la efectividad y el procesamiento de la retroalimentación.....	63
3. USO Y PROCESAMIENTO DE LA RETROALIMENTACIÓN FORMATIVA PARA INCREMENTAR LA COMPRENSIÓN Y APRENDIZAJE DE CONOCIMIENTO CONCEPTUAL.....	68
3.1. Objetivo general.....	68
3.2. Objetivos específicos.....	71
CHAPTER 2: STUDY 1. DOES COMPUTER-BASED ELABORATED FEEDBACK INFLUENCE THE STUDENTS' QUESTION-ANSWERING PROCESS?.....	73
ABSTRACT.....	75
1. INTRODUCTION.....	78
1.1. Question-answering tasks.....	79
1.2. The role of formative feedback in question-answering tasks.....	81
1.3. Objectives and hypotheses.....	84
2. METHOD.....	84
2.1. Participants.....	84
2.2. Instruments.....	84
2.3. Procedure.....	90
2.4. Data analysis.....	90

3. RESULTS.....	91
3.1. Effects of EF on the question-answering performance and the accuracy to assess textual relevance.....	91
3.2. Effects of EF on the question, the text, and the feedback processing.....	92
4. DISCUSSION AND CONCLUSION.....	93
References.....	97

CHAPTER 3: STUDY 2. HOW DO STUDENTS PROCESS COMPLEX FORMATIVE FEEDBACK IN QUESTION-ANSWERING TASKS? A THINK-ALLOUD STUDY.....

ABSTRACT.....	103
1. INTRODUCTION.....	106
1.1. Feedback processing.....	106
1.2. Feedback processing during question-answering tasks.....	108
1.3. Individual differences.....	110
1.4. The current study.....	111
2. METHOD.....	112
2.1. Participants and design.....	112
2.2. Materials.....	113
2.2.1. Texts and questions.....	113
2.2.2. Apparatus.....	114
2.3. Procedure.....	119
2.4. Processing measures.....	120
2.4.1. Think-aloud protocols.....	120
2.4.2. Read&Learn - Behavioral measures.....	123
3. RESULTS.....	124
3.1. Focus of attention on feedback information.....	124
3.2. Attention on feedback conditional to question-answer (success vs. failure)....	125
3.3. Prevalence of cognitive and metacognitive processes during feedback processing and relation to performance.....	126
4. DISCUSSION.....	130
References.....	136

CAPÍTULO 4: ESTUDIO 3: ¿INFLUYE LA RETROALIMENTACIÓN CORRECTIVA EN EL USO DE LA RETROALIMENTACIÓN ELABORADA EN UN ENTORNO DIGITAL?	145
RESUMEN/ABSTRACT.....	147
EXTENDED SUMMARY.....	148
1. INTRODUCCIÓN.....	151
2. MÉTODO.....	156
2.1. Participantes.....	156
2.2. Materiales.....	156
2.3. Herramienta.....	158
2.4. Procedimiento.....	162
2.5. Medidas.....	163
3. RESULTADOS.....	164
3.1. Efecto del tipo de retroalimentación correctiva en el uso de la EF.....	164
3.2. Efecto del tipo de retroalimentación en el rendimiento de los estudiantes.....	166
4. DISCUSIÓN.....	167
Referencias.....	172
CHAPTER 5: STUDY 4: DO SECONDARY-SCHOOL STUDENTS USE ELABORATED FEEDBACK? THE ROLE OF KNOWLEDGE-OF-RESPONSE FEEDBACK, TEXT AVAILABILITY AND PRIOR KNOWLEDGE	177
ABSTRACT.....	179
1. INTRODUCTION.....	180
1.1. Feedback effectiveness and use.....	181
1.2. Text availability.....	183
1.3. Prior knowledge.....	184
1.4. The current study.....	184
2. METHOD.....	185
2.1. Participants and Design.....	185
2.2. Materials.....	186
2.3. Apparatus.....	188
2.4. Procedure.....	192
2.5. Online measures.....	193
3. RESULTS.....	194

3.1. Effects of KR feedback and text availability on accessing EF.....	194
3.2. Effects of KR feedback and text availability on student's performance.....	196
3.3. Does EF access moderate the effect of prior knowledge on student's performance?.....	197
4. DISCUSSION.....	198
References.....	203
CAPÍTULO 6: CONCLUSIONES GENERALES.....	209
REFERENCIAS BIBLIOGRÁFICAS.....	231
ANEXOS.....	259

INTRODUCTION

Students from all over the world are expected to develop reading skills and acquire knowledge as part of the compulsory education. One of the main reasons relies on the fact that students need to use written documents to perform both academic (e.g., answer comprehension questions or write a summary) and non-academic (e.g., read the instructions to fly a drone or read a blog to unlock a character from a videogame) activities. According to the Organization for Economic Co-operation and Development's framework for reading literacy (OECD, 2010), a skilled reader should be able to understand and use a wide variety of texts to achieve different goals. Any of the aforementioned situations involve reading in a very specific way; that is, students have to use the text(s) to do something, so that most reading experiences are driven by goal-directed activities. Therefore, those reading situations can be envisaged as hybrid activities since students need not only to comprehend text information, but also to deploy a number of decisions while performing the tasks (Rouet, Britt, & Durick, 2017).

In school settings, one of the most frequent tasks teachers assign to the students refers to answering comprehension questions from an available text (Ness, 2011). Thus, students are expected to use the document(s) strategically to provide the correct answers (Vidal-Abarca, Mañá, & Gil, 2010), which involves both cognitive (e.g., making inferences) and metacognitive (e.g., assessing one's own understanding, deciding to search the text, or assessing textual relevance) processes. Teachers assign those tasks to assess and/or improve students' comprehension and learning. However, students sometimes struggle to comprehend or use written documents, so that effective instructional procedures need to be designed and applied to overcome those issues. To

that end, teachers (or computers in digital environments) may deliver formative feedback aimed at improving students' abilities or knowledge (Shute, 2008). Here, we focus on question-answering tasks in which formative feedback can be timely delivered in digital learning environments.

Even though the study of feedback has a long history in the field of learning and instruction, only a few studies have attempted to investigate the effects of formative feedback on students' text comprehension and learning from texts. Likewise, research has seldom questioned students' ability and willingness to engage in processing feedback, as well as how it exerts an impact on learning as a function of individual characteristics such as reading skill or prior knowledge. Those are the main goals of the present work. Thus, this thesis builds on previous research conducted in the domains of psychology and education on the feedback effectiveness and the feedback processing. Previous contributions aimed at examining the effects of different types of feedback, including verification information on the students' response (Knowledge of Response or KR), information pointing out the correct answer (Knowledge of Correct Response or KCR), or even more detailed information (Elaborated Feedback or EF). Whereas both KR and KCR include corrective feedback information, the EF encompasses any additional information such as explanations, prompts, or examples. Those feedback messages have been tested on different learning scenarios varying in the nature of the task, ranging from simple memorization or associative tasks (e.g., learning vocabulary or definitions) to higher-order learning tasks (e.g., text comprehension involving inference processes). We propose a framework for both question-answering and feedback processing in the context of question-answering settings based on previous theoretical approaches and empirical findings. Our approach places emphasis on the importance of student's engagement in processing the feedback information actively. Therefore, we posit that feedback is not processed automatically and students have to make decisions consciously about what information to read and how much effort to invest in order to benefit from feedback. Therefore, this thesis includes four studies aimed at examining the processing and effectiveness of various types of EF, taking into account the level of reading skill and prior knowledge of Secondary-school students.

Chapter 1 presents the theoretical framework for both pillars in which this work is framed: question-answering tasks and formative feedback. First, we address the issue of answering questions from an available text, which is one of the most common task-oriented reading activities that school students perform as part of their academic routine

(Ness, 2011). In this section, we present a model of question-answering tasks to point out the processes involved since these environments require students not only to comprehend information, but also to make decisions on the fly as they perform the task. We then address the individual differences found in question-answering tasks. Second, we review previous research on formative feedback, emphasizing its effects on text comprehension and learning conceptual knowledge. In question-answering tasks, teachers may provide students with formative feedback based on their performance outcomes. However, computer-based systems are able to deliver more specific and tailored feedback messages. Therefore, we mainly review previous evidence of the effects and the use of different types of feedback within the context of digital environments with the aim of developing a framework for feedback processing that encompasses the cognitive and metacognitive processes involved. Finally, we analyze the potential effects of formative feedback as a function of individual differences. The last section of this chapter covers the general and specific goals of the thesis in light of the theoretical review carried out and the need to understand how computer-based feedback is processed.

In Chapters 2, 3, 4, and 5, we present four studies to better understand how students interact with computer-based feedback in question-answering settings and its effects on text comprehension and learning outcomes while taking into account the students' reading skill and prior knowledge. Each of the studies is presented in the format of a scientific article, including a theoretical introduction, the method employed, the results found and the corresponding discussion. Chapter 2 presents the first study in which we examined the influence of a detailed EF in both question-answering performance and accuracy to assess textual relevance of high school students. Further, we analyze to what extent that EF was processed over a control condition in which non-formative feedback was delivered. Likewise, we explored whether reading skill influenced how students engaged in processing EF differently. 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). Main findings

suggested that EF influenced positively the students' text comprehension performance and the assessment of textual relevance (i.e., EF reduced the amount of non-relevant text information students assessed as question-relevant). However, EF did not affect neither the question nor the text-search processing times, suggesting that EF improved students' efficiency to search and assess textual relevance. Regarding feedback processing, findings showed that EF increased the processing times and the decision-making process to access optionally-delivered feedback information above the control condition. Additionally, we found that skilled comprehenders outperformed their less-skilled counterparts. This study sheds light on how a complex EF on the students' question-answering performance and their accuracy to assess textual relevance may improve the students' question-answering process in a digital environment. However, further research is necessary to explore the cognitive and metacognitive processes involved in feedback processing, as well as how skilled and less-skilled comprehenders engage in processing such an EF.

In Chapter 3, we explicitly seek to explore what components of EF students pay attention to and what cognitive and metacognitive processes students deploy when EF is delivered in a digital environment. According to the results found in the first study, we conducted a second study (Study 2) with the same materials. In this case, however, skilled and less-skilled comprehenders in grade 8 were asked to report any thought that came to their minds while answering the questions and receiving EF. Participants thought-aloud in one text and performed the task in silence on the other text. Main findings showed that students paid more attention to know whether their answers were correct than to any other elaborative components of feedback. Related to the previous result, we found that students focused their attention on the incorrectly-answered questions, but paid little or no attention to feedback on those questions they succeeded. Whereas students actively monitored the accuracy of their responses by comparing their answers with the standard provided, they rarely constructed meaning and self-regulated the use of feedback. Interestingly, feedback triggered affective reactions and attributions as part of the monitoring processes. Individual differences in feedback processing suggested that skilled and less-skilled comprehenders processed the EF quite similarly. Those findings suggested that other individual characteristics like prior knowledge may be more relevant to understand how students engage in and process EF in the context of learning complex knowledge from academic texts.

Chapter 4 presents the study 3. Based on previous findings, we changed the approach of feedback processing and learning from texts due to the fact that students were mainly interested in knowing the correctness of their responses or the correct response rather than engaging in deep processing. Maybe students processed EF narrowly in the previous studies because the EF messages included the correct response (i.e., KCR feedback), which may have discouraged students to process additional information. Likewise, we employed a question-answering task designed to assess student's text comprehension, not learning from texts. Hence, we designed an experiment to study how the presence of corrective feedback (KCR, KR, or control feedback) may influence the student's decision to use additional EF made of explanations when learning conceptual knowledge in the area of physics. Due to the topic covered in the science text (*'Atmospheric pressure and the wind phenomenon'*), we took into account the student's prior knowledge, arguably one of the main individual factors involved in text comprehension and learning from texts. Secondary-school students in grade 9 answered a set of questions from an available science text. For each question, they received corrective feedback according to the condition assigned: KCR, KR, or Control (i.e., non-corrective feedback). After receiving this feedback, all the students were allowed to access item-based EF that included an explanation about the knowledge assessed. 24 hours later, students completed a final test with new questions. Main findings showed that students do not use EF very often, especially when corrective feedback includes the correct response (i.e., KCR feedback). Further, results showed that KR and KCR feedback made students focus on EF for incorrectly-answered questions. Finally, results suggested that no differences were found for the learning task and the final performance outcomes.

In Chapter 5, we present study 4, aimed at examining how two external factors (corrective feedback and text availability) may influence Secondary-school 9 grade students' decision to use additional EF and its consequent impact on learning science knowledge about the atmospheric pressure and the wind phenomenon. Based on the previous findings, we decided to run the experiment without the KCR feedback condition because students decided to access little EF messages when this corrective feedback was delivered. Thus, we included the KR and the control feedback groups in this experiment. Additionally, we manipulated the text availability while answering the questions, so that one half of the sample performed the task with text access and the other half performed the task without text access. Main findings showed that corrective

feedback made students focus on EF for incorrectly-answered questions, and that keeping the text unavailable made students use EF more often. However, neither corrective feedback nor text availability influenced student's learning outcomes. Moreover, the relation between students' prior knowledge and learning task performance was moderated by the students' decision to access EF, which means that students with lower levels of prior knowledge who accessed EF quite often reduced the impact of their prior knowledge. Nevertheless, both prior knowledge and EF accesses played independent positive roles in the final task administered several hours later.

Finally, Chapter 6 presents the general conclusions and the limitations of the studies conducted, as well as the theoretical and practical implications for both researchers and practitioners in the fields of psychology and education interested in enhancing students' learning by means of delivering formative feedback in digital environments. Hence, we point out the importance of understanding students' willingness to make use of formative feedback and, as a consequence, the need to develop digital environments where students have the opportunity to use feedback on demand, since automatic feedback delivery does not guarantee that students are driven to attend and process its content in the expected manner. Thus, we provide educational responses about how to deliver computer-based feedback to enhance students' text comprehension and learning in digital platforms like e-textbooks, MOOCs (Massive Online Open Courses), or ITSs (Intelligent Tutoring Systems). We will conclude the thesis with some possible paths for future research on computer-based feedback processing and effectiveness in the context of question-answering scenarios.

CAPÍTULO 1

MARCO TEÓRICO

1. EL PAPEL DE LAS PREGUNTAS PARA EVALUAR LA COMPRESIÓN Y EL APRENDIZAJE

A lo largo de los últimos años ha habido un creciente interés por comprender la lectura y comprensión de textos en situaciones reales en las que los lectores leen documentos con un propósito determinado (e.g., Snow & the RAND Reading Study Group, 2002). Mientras que los estudios sobre comprensión lectora de las últimas décadas se han centrado en estudiar los procesos de comprensión implicados en la construcción de representaciones mentales de un texto (e.g., realización de inferencias), las perspectivas actuales amplían esta visión y ponen especial énfasis en el uso de los textos para alcanzar objetivos específicos. Según el reciente modelo RESOLV (REading as problem SOLVing; Britt, Rouet y Durik, 2018), la lectura puede conceptualizarse como una actividad dirigida a resolver problemas, por lo que los lectores han de usar los textos con el objetivo de completar tareas específicas como, por ejemplo, responder preguntas o hacer un esquema. Estas situaciones se corresponden con la actual perspectiva teórica de competencia lectora, desde la que se conceptualiza al lector competente como aquel que tiene capacidad para utilizar la información textual para alcanzar objetivos específicos (OECD, 2017). Así, el lector es un buscador activo de información que adapta la lectura para lograr esos objetivos (Britt et al., 2018).

En el contexto académico, los estudiantes participan diariamente en una amplia variedad de actividades que implican la lectura de textos con diferentes propósitos. Habitualmente, los docentes asignan tareas a sus alumnos (e.g., contestar preguntas, hacer un esquema, escribir un resumen) con la finalidad de facilitar la adquisición de conocimientos y el desarrollo de habilidades (National Reading Panel, 2000). Así, los

estudiantes suelen realizar lecturas y resolver tareas diseñadas para aprender conocimientos como, por ejemplo, cuáles son las causas y consecuencias de la Revolución Francesa, o cómo la temperatura y la altura influyen en la presión atmosférica. En estas tareas, los estudiantes conocen de antemano que los textos serán una fuente esencial de información que estará disponible cuando resuelvan la tarea, por lo que podrán consultar los materiales para buscar información relevante cuando lo consideren oportuno (Vidal-Abarca et al., 2010).

Estudios recientes demuestran que las actividades de lectura más frecuentes en las aulas consisten en leer un texto y posteriormente evaluar la comprensión de los alumnos mediante preguntas sobre sus contenidos (Sánchez y García, 2015; Sánchez, García y Rosales, 2010). Éste es el tipo de tareas en el que nos centraremos en esta tesis. Frecuentemente, los estudiantes responden preguntas de comprensión con acceso a los materiales, una actividad que respalda el aprendizaje asociado con la lectura (Anmarkrud, McCrudden, Bråten y Strømsø, 2013; Cerdán, Gilabert y Vidal-Abarca, 2011; Golding, Graesser y Millis, 1990; Graesser y Murachver, 1985; Graesser, Robertson, Lovelace y Swinehart, 1980; Lewis y Mensink, 2012; Mañá, Vidal-Abarca y Salmerón, 2017; Roelle y Berthold, 2017; Vidal-Abarca et al., 2010). En general, podemos decir que responder preguntas a partir de textos permite a los estudiantes aprender conceptos e ideas, ya que facilita la integración de información nueva en la red de conocimientos previamente adquiridos, debido al papel que juegan las preguntas en el procesamiento de los contenidos del texto (e.g., Burton y Daneman, 2007; Kaakinen, Hyönä y Keenan, 2003), ya que buscar información para responder preguntas supone evaluar la relevancia de la información textual con el objetivo de localizar información pertinente para la tarea (McCrudden y Schraw, 2007). A pesar de la frecuencia de este tipo de actividades en el contexto educativo, existe poca investigación que haya explorado los procesos involucrados al responder preguntas de un texto disponible (e.g., Ferrer, Vidal-Abarca, Serrano y Gilabert, 2017; Ozuru, Best, Bell, Witherspoon y McNamara, 2007; Schaffner y Schiefele, 2013; Schroeder, 2011).

Estudios observacionales en entornos escolares ponen de manifiesto que los profesores asignan tareas consistentes en responder preguntas de los libros de texto para apoyar el aprendizaje de sus alumnos (Durkin, 1978-1979; Ness, 2011; Pressley, Wharton-McDonald, Mistretta-Hampston y Echevarria, 1998). En un estudio con alumnos de primaria, Ness (2011) encontró que la estrategia de comprensión lectora más empleada por los docentes consistía en pedir a los estudiantes que respondieran

preguntas de un texto. En estas situaciones, los estudiantes utilizaban el texto para buscar las respuestas, y posteriormente los profesores les proporcionaban retroalimentación sobre la precisión de las mismas. Este tipo de actividades en las que los estudiantes han de resolver tareas a partir de textos para comprender/aprender sus contenidos implica que los aprendices monitoricen su progreso y autorregulen su aprendizaje activamente (Pintrich, 2000; Winne, 2004).

En entornos digitales de aprendizaje, las tareas suelen tener una estructura similar; es decir, los estudiantes leen (o visionan) uno o varios documentos, responden una serie de cuestiones sobre los materiales, y reciben retroalimentación sobre su rendimiento (Conole, 2016). Cada vez es más frecuente ver cómo los alumnos completan estas tareas en dispositivos tecnológicos como ordenadores o tablets, debido en parte al incremento en el número de dispositivos tecnológicos en los centros educativos de los países europeos (e.g., European Commission, 2013). En este sentido, los sistemas de tutorización inteligente (*Intelligent Tutoring Systems* o *ITSs*) abren nuevos caminos a la aplicación de este tipo de tareas junto con procedimientos instruccionales que faciliten el aprendizaje de los estudiantes como, por ejemplo, la retroalimentación formativa (Graesser, 2016; Nye, Graesser y Hu, 2014). Aunque estos entornos digitales ofrecen una serie de funcionalidades destacables (e.g., corrección de respuestas automáticamente y entrega de retroalimentación de forma inmediata), es necesario comprender cómo utilizan (cognitiva y metacognitivamente) la retroalimentación formativa los estudiantes de secundaria al responder preguntas de un texto en un entorno digital. Éste es el objetivo central de este trabajo. Además, dado que las diferencias individuales basadas en el nivel de competencia lectora y el conocimiento previo están bien documentadas en este tipo de tareas de aprendizaje, es importante comprender su papel en el uso y procesamiento de la retroalimentación formativa.

1.1. Marco de trabajo para las tareas de responder preguntas

Los libros de texto suelen proponer tareas tales como responder preguntas para facilitar la comprensión y aprendizaje de determinados conceptos e ideas. Habitualmente, los estudiantes resuelven estas tareas sabiendo de antemano que el texto será un recurso disponible donde localizar información relevante para las mismas (Anmarkrud et al., 2013; Gil, Martínez y Vidal-Abarca, 2015; Vidal-Abarca et al.,

2010). Resolver tareas de aprendizaje a partir de uno o varios textos requiere que los estudiantes apliquen procesos cognitivos y metacognitivos continuamente, ya que han de interactuar con los materiales (i.e., textos y tareas) (Rouet, 2006; Rouet y Britt, 2011) para localizar la información relevante para la tarea (McCrudden, Magliano y Schraw, 2011; McCrudden y Schraw, 2007). Cada tarea puede concebirse como un conjunto de instrucciones que modifican los procesos de lectura y los resultados de aprendizaje, ya que dirige la atención de los estudiantes hacia la información relevante (i.e., información percibida como útil para resolver una tarea específica). Este fenómeno se conoce como el *efecto de relevancia* (McCrudden y Schraw, 2007), según el cual los lectores tienden a procesar más detenidamente la información relevante para la tarea en comparación con la información menos relevante.

Hasta la fecha, existen varias investigaciones que han examinado los procesos cognitivos y metacognitivos involucrados en responder preguntas a partir de textos (e.g., Cataldo y Oakhill, 2000; Cerdán et al., 2011; Cerdán, Vidal-Abarca, Martínez, Gilabert y Gil, 2009; Vidal-Abarca, Salmerón y Mañá, 2011). Según el modelo RESOLV (Rouet et al., 2017), responder preguntas de un texto disponible es una actividad híbrida que implica tanto procesos de comprensión como de resolución de problemas, ya que la lectura se dirige al objetivo específico de dar la respuesta correcta a cada pregunta (e.g., Farr, Pritchard y Smitten, 1990; Rupp, Ferne y Choi, 2006). En primer lugar, los procesos básicos de comprensión (e.g., comprensión de ideas explícitas, generación de inferencias entre ideas del texto, activación de conocimientos previos, elaboración de macro-ideas) conforman un elemento esencial en estas tareas, ya que los estudiantes han de construir una representación mental del contenido del texto (Graesser, Singer y Trabasso, 1994; Kintsch, 1998; Kintsch y van Dijk, 1978; van Dijk y Kintsch, 1983). En segundo lugar encontramos las estrategias metacognitivas, las cuales juegan un papel esencial en la resolución de las tareas, ya que el estudiante necesita tomar decisiones continuamente para alcanzar sus objetivos (Rouet, 2006; Rouet y Britt, 2011; Rouet et al., 2017). Estas decisiones se relacionan con el manejo de la información textual para cada pregunta en concreto. El estudiante debe tomar decisiones sobre cuándo buscar información relevante, dónde localizarla, cómo procesarla, o cuándo detener la búsqueda cuando ha localizado la información necesaria (Vidal-Abarca et al., 2010). Este conjunto de decisiones juega un papel importante en el rendimiento de los estudiantes (Gil et al., 2015).

El modelo RESOLV también asume que los estudiantes forman un modelo de tarea y un modelo del texto. El modelo de tarea consiste en una representación mental de cada pregunta que incluye el objetivo del estudiante y un conjunto de medios para alcanzarlo (e.g., búsquedas estratégicas de información en el texto). En el caso de responder varias preguntas, el estudiante construirá un modelo de tarea para cada pregunta. El modelo del texto, por su parte, es una representación mental sobre el texto que ha de estar alineada con los objetivos de aprendizaje del estudiante. La formación de estas representaciones requiere tanto procesos cognitivos que apoyen la respuesta a las preguntas (e.g., Cerdán et al., 2011; Graesser y Clark, 1985; Graesser y Franklin, 1990) como procesos metacognitivos para responderlas (e.g., decidir cuándo buscar información en el texto o monitorizar qué información es relevante para la tarea) (Mañá et al., 2017; Vidal-Abarca et al., 2010). Estos procesos son diferentes de los que podemos encontrar en situaciones de comprensión lectora tradicionales (Cataldo y Oakhill, 2000; Rouet, 2006), ya que responder preguntas con acceso a los textos engloba un conjunto de estrategias metacognitivas propias que no están presentes cuando las tareas se realizan sin acceso a los textos (e.g., Artelt, Schiefele y Schneider, 2001; Ferrer et al., 2017; Ozuru et al., 2007). Esta cuestión es esencial para los estudios presentados a lo largo de esta tesis, ya que el cuarto estudio introduce la manipulación de la disponibilidad de los textos para examinar su impacto en el uso de la retroalimentación. En ese caso, disponer o no de la información textual durante el proceso de responder preguntas elimina la posibilidad de buscar información relevante para la tarea, por lo que la disponibilidad de los materiales puede jugar un papel importante en la decisión de utilizar otros recursos de información disponibles (e.g., retroalimentación formativa que incluya explicaciones sobre los contenidos evaluados en cada pregunta). A continuación se presenta un modelo que representa las principales acciones y decisiones que puede realizar un estudiante durante el proceso de responder preguntas de un texto disponible seguido de retroalimentación externa, el otro pilar sobre el que se fundamentan los estudios de este trabajo (Figura 1). Este modelo se basa en resultados de investigaciones previas en tareas de contestación a preguntas (e.g., Cerdán et al., 2009; Gil et al., 2015; Vidal-Abarca et al., 2010) y propuestas teóricas como el modelo TRACE (*Task-based Relevance and Content Extraction*; Rouet, 2006) o el *Goal-focusing model* (McCrudden y Schraw, 2007).

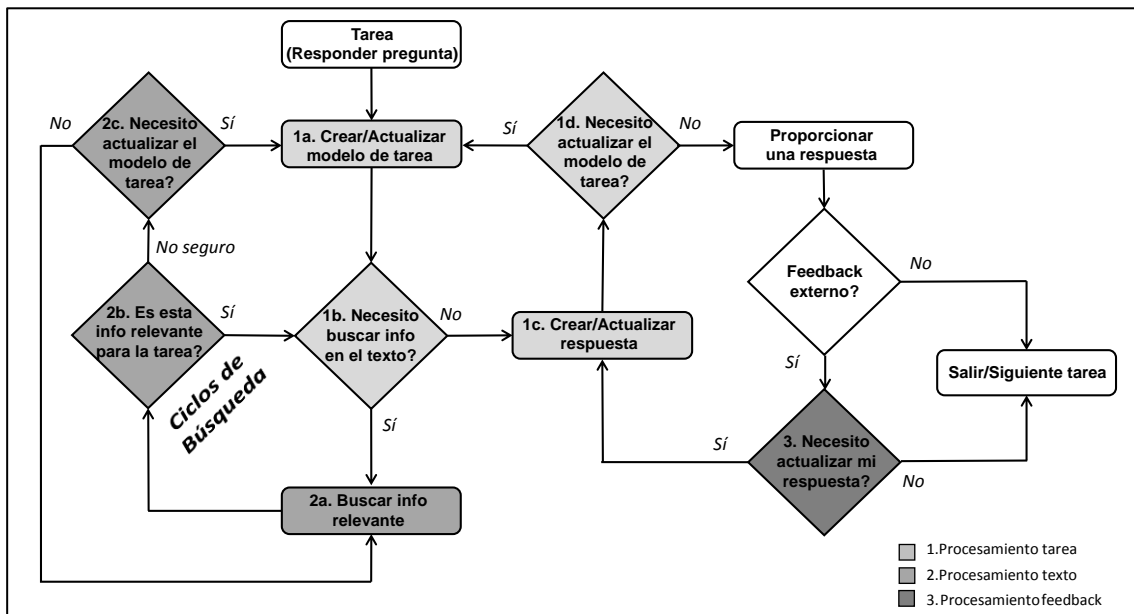


Figura 1. Modelo QAL (Question-Answering for Learning) tomado y traducido de Vidal-Abarca, Martínez, Gil, García y Máñez (2019).

Realizar tareas de aprendizaje con acceso a los textos implica tomar decisiones durante el procesamiento de la tarea, la búsqueda de información relevante en el texto, y el procesamiento de la retroalimentación en caso de que se proporcione. Durante la resolución de las tareas, éstas deben ser interpretadas por el alumno, quien debe construir un modelo de tarea (1a), es decir, una representación mental que incluya los objetivos de la tarea y los medios disponibles para alcanzarlos (Rouet et al., 2017). Una vez que se ha creado el modelo de tarea, el estudiante ha de tomar un conjunto de decisiones acerca de cómo procesar el texto. Así, el estudiante ha de decidir si necesita o no buscar información en el texto para crear una respuesta (1b), lo que implica monitorizar los procesos que operan en el modelo de tarea previamente construido. Si el estudiante considera que puede responder la pregunta con la información que tiene activa en su memoria de trabajo sin necesidad de consultar el texto, entonces creará y dará una respuesta con la información que recuerde (1c). Por el contrario, si el estudiante considera que necesita consultar el texto, entonces iniciará un nuevo proceso de localización de información relevante para contestar la pregunta (2a). En este momento, el estudiante ha de evaluar qué información es relevante (2b) para alcanzar sus objetivos (McCrudden y Schraw, 2007) y responder a las demandas de la tarea (Rouet, 2006). Esta búsqueda de información opera sobre el modelo de tarea construido anteriormente y funciona en ciclos de interacción Tarea-Texto hasta que el estudiante

considera que ha alcanzado el objetivo propuesto o que no puede mejorar su respuesta. A lo largo del proceso de resolver una pregunta, podemos asumir que la construcción del modelo de tarea es un proceso dinámico que puede durar hasta el momento de proporcionar una respuesta, ya que el alumno puede modificar este modelo continuamente (Rouet et al., 2017). Durante la búsqueda de información, el estudiante puede modificar su modelo de tarea en función de la información textual procesada (2c). Asimismo, durante el proceso de creación/actualización de la respuesta, el estudiante también podrá actualizar su modelo de tarea (1d). A lo largo del proceso de responder las preguntas, los estudiantes evalúan la consecución de sus objetivos, ajustando así sus conductas a las demandas de la tarea. Al responder, es posible que el estudiante reciba retroalimentación externa. Esta información, al ser procesada, permite al estudiante evaluar la necesidad de actualizar su modelo de respuesta. Si esto sucede, puede que se produzcan modificaciones en la representación construida sobre el contenido del texto, bien añadiendo nueva información, bien re-estructurando los conocimientos previamente aprendidos (3).

Como se ha mencionado anteriormente, el modelo de tarea es una interpretación que realiza el estudiante sobre el contenido de la pregunta y que puede ser modificada durante la resolución de la misma (1a). La precisión con la que el estudiante construya el modelo de tarea afectará al rendimiento que obtenga (Rouet, 2006; Rouet y Britt, 2011). Si el estudiante es capaz de comprender la pregunta y sabe gestionar el uso de los recursos disponibles, la probabilidad de proporcionar la respuesta correcta se incrementará. Formar el modelo de tarea implica tanto procesos cognitivos (e.g., formar e integrar las diferentes proposiciones que componen la pregunta) como procesos metacognitivos (e.g., evaluar la comprensión y la memoria de la información aprendida), lo que determinará las decisiones de búsqueda (i.e., qué información buscar y dónde buscarla). Durante la construcción del modelo de tarea, el estudiante ha de monitorizar el grado de comprensión de la información leída. Una técnica efectiva para conocer si un estudiante monitoriza adecuadamente la comprensión de la información es el paradigma de detección de errores, consistente en introducir una contradicción en la información al objeto de examinar si el estudiante es capaz de identificarla. En el caso de tareas consistentes en responder preguntas, resultados recientes sugieren que la capacidad de monitorizar la comprensión de las preguntas (i.e., determinar si una pregunta contiene una contradicción) predice significativamente el rendimiento en la tarea (Mañá, Vidal-Abarca, Domínguez, Gil y Cerdán, 2009).

Asimismo, existen evidencias de que el nivel de procesamiento necesario para responder preguntas varía según la complejidad para construir el modelo de tarea. Cerdán y Vidal-Abarca (2008) y Cerdán et al. (2009) examinaron el procesamiento de la tarea distinguiendo entre preguntas de bajo o alto nivel. Por ejemplo, Cerdán et al. (2009) pidieron a los estudiantes que leyeran un texto de 2000 palabras sobre modelos atómicos y luego respondieran preguntas de bajo o alto nivel según la condición. Los participantes que contestaron preguntas de bajo nivel respondieron 19 preguntas que implicaban la identificación de segmentos específicos de información, mientras que los participantes que contestaron preguntas de alto nivel respondieron 5 preguntas que implicaban integrar información de diferentes párrafos. Ambos tipos de preguntas cubrían la misma información textual, por lo que ambas condiciones abarcaban la misma información relevante. Sin embargo, las operaciones mentales necesarias para responder fueron diferentes. Mientras que las preguntas de bajo nivel requerían que el estudiante localizara la respuesta en segmentos específicos del texto mediante la realización de inferencias sencillas, las preguntas de alto nivel requerían que el estudiante integrara información de varios párrafos mediante inferencias más complejas entre segmentos distantes o con su conocimiento previo. Dos días más tarde, los participantes completaron una tarea de recuerdo y contestaron una serie de preguntas que implicaban inferencias complejas. Los resultados mostraron que los estudiantes que respondieron preguntas de bajo nivel obtuvieron un rendimiento más alto que aquellos estudiantes que respondieron preguntas de alto nivel. En el post-test, aunque no hubo diferencias en el rendimiento en la tarea de recuerdo, los estudiantes que respondieron preguntas de alto nivel obtuvieron un mejor rendimiento en las preguntas inferenciales que aquellos que respondieron preguntas de bajo nivel durante la fase de aprendizaje. Por tanto, los autores concluyeron que contestar preguntas de alto nivel supone mejoras en términos de aprendizaje profundo en comparación con las de bajo nivel. Además de los resultados de rendimiento, Cerdán et al. (2009) encontraron que contestar preguntas de alto nivel afecta al procesamiento de la información, ya que los estudiantes leyeron las preguntas más detenidamente, relevaron más información relevante, y realizaron más interacciones entre la pregunta, el texto, y la redacción de sus respuestas. En una tarea similar con documentos múltiples, Cerdán y Vidal-Abarca (2008) encontraron que contestar una pregunta de alto nivel que requería integrar información relevante de tres textos redujo el procesamiento de segmentos aislados de información relevante y mejoró el aprendizaje profundo en comparación con contestar preguntas de bajo nivel.

Una vez que el alumno ha formado el modelo de tarea, éste ha de decidir si necesita consultar el texto para buscar información relevante que le ayude a responder correctamente (2a). Por tanto, para activar este proceso de autorregulación, los estudiantes deben monitorizar su nivel de comprensión y evaluar si serían capaces de responder correctamente con la representación mental construida. Si el estudiante considera que no tiene suficiente conocimiento para responder la pregunta (i.e., no puede activar en su memoria de trabajo la información relevante para la tarea en cuestión), entonces tal vez inicie un proceso de autorregulación consistente en buscar estratégicamente la información relevante en el texto. Esta decisión depende de habilidades metacognitivas y juega un papel esencial en el proceso de responder preguntas de un texto (Gil et al., 2015; Mañá et al., 2017; Rouet y Coutelet, 2008). En el ámbito del aprendizaje autorregulado, la monitorización es un componente crucial que determina las decisiones que toman los estudiantes, por lo que las dificultades para monitorizar el grado de aprendizaje repercuten negativamente en la eficacia para autorregular el propio proceso de aprendizaje y el consiguiente rendimiento final (Metcalf, 2002; Metcalf y Finn, 2008; Thiede, Anderson y Therriault, 2003; Thiede y Dunlosky, 1999).

Por lo que respecta a la monitorización de la búsqueda de información al responder preguntas de un texto, Vidal-Abarca et al. (2010, experimento 2) examinaron si las decisiones de consultar el texto son independientes de las habilidades de comprensión de los estudiantes. Para ello pidieron a estudiantes de secundaria que contestaran una serie de preguntas en un ordenador al objeto de registrar las decisiones de consultar el texto al contestar las preguntas. Después de leer cada pregunta, los participantes debían indicar en qué medida podían contestar correctamente la pregunta sin consultar el texto en una escala de 0 (*Seguro que no podré*) a 100 (*Seguro que podré*) con intervalos de 20 en 20. Los resultados mostraron que los estudiantes consultaron el texto cuando no estaban seguros de poder responder la pregunta sin buscar información, mientras que decidieron no consultar el texto cuando estaban seguros de responder correctamente sin buscar información. Este resultado sugiere que la decisión de buscar información en el texto depende de procesos de monitorización acerca de la necesidad de localizar información relevante para lograr sus objetivos. Además, Vidal-Abarca et al. (2010) encontraron que la precisión de la monitorización es independiente de las habilidades de comprensión de los estudiantes. Por su parte, Gil et al. (2015) encontraron que la decisión de no buscar información cuando es

innecesario y la cantidad de decisiones de buscar información predijeron significativamente el rendimiento en tareas de competencia lectora. Adicionalmente, Mañá et al. (2009) encontraron que los estudiantes experimentan dificultades a la hora de monitorizar la decisión de no consultar el texto al responder preguntas sobre los contenidos del mismo. Los resultados revelaron que los estudiantes respondían correctamente aproximadamente la mitad de las preguntas cuando decidían no buscar información en el texto y confiaban en su representación mental.

Una vez que el estudiante ha tomado la decisión de buscar información en el texto, inicia el proceso de búsqueda de información relevante que responda a las demandas del modelo de tarea construido (2b). Así, el estudiante ha de identificar qué información del texto es relevante para resolver la tarea y debe tomar decisiones acerca de cuándo detener su búsqueda. Para que la búsqueda resulte efectiva, el estudiante ha de aplicar procesos de comprensión que le permitan conectar información textual con el modelo de tarea que mantiene activo en su memoria de trabajo para proporcionar una respuesta. Para ello, el estudiante ha de monitorizar la búsqueda, evaluando el grado de relevancia de la información que decide leer, al objeto de seleccionar segmentos de información relevante y descartar aquellos segmentos con información menos relevante o irrelevante. Por tanto, este procesamiento de la información textual orientado a la contestación de preguntas requiere de habilidades específicas que van más allá de los procesos básicos de comprensión (e.g., Cataldo y Oakhill, 2000; Mañá et al., 2009; Rouet, 2006).

Investigaciones recientes han demostrado que la precisión al identificar información relevante durante la búsqueda explica el rendimiento de los estudiantes (Gil et al., 2015; Mañá et al., 2009). De hecho, se han encontrado relaciones entre conductas estratégicas de búsqueda de información y el rendimiento al responder preguntas (e.g., Cerdán et al., 2009; Gil et al., 2015). Por ejemplo, los estudiantes que responden correctamente más preguntas tienden a generar sus respuestas inmediatamente después de procesar información relevante para la tarea (Cerdán et al., 2009). Es decir, son capaces de discriminar información relevante y utilizarla para proporcionar respuestas. Cerdán et al. (2011) encontraron que aquellos estudiantes que emplean estrategias de análisis semántico al buscar información relevante obtienen mejores resultados que aquellos que se guían por estrategias superficiales basadas en la superposición de palabras entre la pregunta y el texto. Aunque todos los estudiantes tienden a utilizar estrategias superficiales de superposición de palabras inicialmente, aquellos con altas

habilidades de comprensión son capaces de modificar su estrategia y realizar un análisis semántico que les permita identificar la verdadera información relevante (Cerdán et al., 2011).

Aunque este modelo no contempla el procesamiento del texto antes de responder las preguntas (i.e., lectura/estudio del texto inicialmente), cabe señalar que esta fase puede influir tanto en el proceso de búsqueda de información durante la tarea como sobre el rendimiento final. La lectura inicial del texto puede ayudar a los estudiantes a construir una representación mental coherente basada en su conocimiento previo (Kintsch, 1998). Sin embargo, el estudiante desconoce qué información textual será relevante para resolver cada una de las tareas/preguntas (McCrudden y Schraw, 2007). Así, los estudiantes han de tomar decisiones acerca de cómo procesar esta información inicialmente (e.g., no leer el texto y buscar información a posteriori, leer sólo los títulos de las secciones para conocer la estructura del texto, leer sólo el comienzo para conocer de qué trata, o leer el texto completo). Investigaciones previas demuestran que los estudiantes tienden a leer el texto de principio a fin antes de pasar a las preguntas (Salmerón, Vidal-Abarca, Mañá, Martínez, Gil y Naumann, 2015), aunque esta lectura inicial suele ser superficial cuando saben que el texto será una fuente de información disponible durante la tarea (e.g., Ferrer et al., 2017; Higgs, Magliano, Vidal-Abarca, Martínez y McNamara, 2017). Si el estudiante decide leer las preguntas en primer lugar, el procesamiento del texto estará guiado por las demandas de cada pregunta, por lo que evaluará directamente la relevancia textual de la información que lea. Resultados recientes sugieren que leer el texto inicialmente favorece el rendimiento al responder preguntas de comprensión (Cerdán et al., 2009; Gil et al., 2015; Salmerón et al., 2015), ya que permite a los estudiantes construir una representación mental del texto que podrá ser utilizada para responder las preguntas (Rouet, 2006) y para buscar información relevante en el texto (Cataldo y Oakhill, 2000; Cerdán et al., 2009; Vidal-Abarca et al., 2010). En los estudios que se presentan en este trabajo nos aseguramos de que los participantes leyeran el texto antes de responder las preguntas con la intención de minimizar los efectos debidos a la variabilidad en la lectura inicial.

Al completar las tareas, los entornos digitales pueden proporcionar retroalimentación externa, cuya función última es modificar los procesos cognitivos y/o comportamentales de los estudiantes para mejorar su aprendizaje (Shute, 2008). Según el informe del National Reading Panel (2000), evaluar la comprensión mediante preguntas y proporcionar retroalimentación inmediata es una de las estrategias eficaces

para mejorar la comprensión de los estudiantes. En las tareas de responder preguntas de un texto, la retroalimentación suele dirigirse a verificar el conocimiento evaluado, corregir conocimientos erróneos, y/o desarrollar habilidades de comprensión. De hecho, el aprendizaje se concibe no sólo como la adquisición de nuevos conocimientos, sino también como la reestructuración de los conocimientos previamente adquiridos (Brod, Werkle-Bergner y Shing, 2013). Si se proporciona retroalimentación externa después de responder una pregunta, los estudiantes pueden utilizar esta información para corroborar sus conocimientos o para modificar su modelo de respuesta, especialmente después de proporcionar respuestas incorrectas (3). En general, cuando se proporciona retroalimentación al responder preguntas de un texto, el objetivo es producir cambios conceptuales en la representación mental del estudiante, por lo que las ideas o relaciones entre ideas erróneas han de ser corregidas o eliminadas. Para que esto tenga lugar, es necesario que los estudiantes estén dispuestos y sean capaces de procesar su contenido activamente (Carless y Boud, 2018; Timmers y Veldkamp, 2011). En caso de que decidan procesarlo, se espera que la retroalimentación permita comparar la ejecución del estudiante con un estándar y así reducir la distancia entre su nivel de rendimiento actual y el deseado. No obstante, no todos los tipos de retroalimentación contienen la misma información, requieren el mismo nivel de procesamiento, ni son igualmente efectivos (Shute, 2008; Van der Kleij, Feskens y Eggen, 2015). En el segundo apartado de la introducción expondremos en detalle los resultados más relevantes de la investigación sobre retroalimentación al realizar tareas de comprensión y aprendizaje, prestando especial atención a la investigación realizada en entornos digitales.

En términos generales, podemos afirmar que responder preguntas de un texto es una actividad común que los docentes suelen emplear para: a) facilitar la comprensión de la información del texto, y b) promover el aprendizaje de sus contenidos. Cuando evalúan el nivel de comprensión, las preguntas rastrean la capacidad del estudiante para localizar información relevante específica, generar inferencias o crear macroideas, con la intención de ayudarlo a construir una representación mental coherente e integrada de la información contenida en el texto. Es decir, las preguntas fomentan la construcción de una representación mental conocida como base del texto. Cada pregunta se puede considerar como un conjunto de instrucciones que guían el procesamiento de la información, dirigiendo la atención hacia aquellos segmentos de información que son relevantes para la tarea (McCrudden y Schraw, 2007). Sin embargo, cuando evalúan el grado de adquisición o aprendizaje de nuevos conocimientos, las preguntas no sólo

evalúan la capacidad del estudiante para comprender la información del texto, sino que además evalúan su capacidad para descontextualizar la información procesada (i.e., evalúan la capacidad para aplicar los conocimientos aprendidos a nuevas situaciones, bien en la misma sesión de aprendizaje, bien en una nueva sesión tras un periodo de demora). Por tanto, estas preguntas evalúan la capacidad para construir un modelo de la situación que vaya más allá de los contenidos del texto. A pesar de estas diferencias, cabe señalar que la comprensión de textos y el aprendizaje a partir de textos guardan una estrecha relación, ya que para utilizar la información en nuevas situaciones (perspectiva del aprendizaje) es necesario integrar la información del texto en nuestro conocimiento previo a fin de construir una representación mental coherente del texto (perspectiva de la comprensión) (Kendeou, Rapp y van den Broek, 2004; van den Broek, 2010).

Esta distinción es importante para comprender el conjunto de estudios que presentamos en este trabajo. Los dos primeros estudios utilizan preguntas de un test estandarizado que fueron diseñadas para examinar el grado de comprensión de textos expositivos no académicos (*Test de Procesos de Comprensión*; Martínez, Vidal-Abarca, Sellés y Gilabert, 2008). Imagínese, por ejemplo, a un estudiante contestando preguntas de un texto sobre ‘Los Sioux’, una tribu de nativos americanos que ocupaban territorios del norte de los Estados Unidos y Canadá. Esas preguntas evalúan la capacidad del estudiante para realizar las operaciones cognitivas propias para comprender los contenidos del texto. Por ejemplo, imagine que ha de contestar la siguiente pregunta: ‘*El gobierno intentó inicialmente encerrar a los sioux en las reservas para que:*’. Para responderla correctamente, el estudiante ha de realizar una inferencia entre estos dos segmentos de información relevante para la pregunta: (a) ‘*Los colonos blancos buscando tierras y los mineros en busca de oro iniciaron una continua invasión de los territorios indios*’, y (b) ‘*El gobierno intentó encerrar a los sioux en reservas*’. Si el estudiante es capaz de realizar la inferencia, entonces comprenderá que el gobierno encerró a los sioux para que ‘*Los colonos y los mineros ocuparan las tierras indias.*’, debido a los intereses por conseguir una mayor extensión de territorio o por hacerse con los recursos de la región. Como se puede apreciar, el principal objetivo es conocer si el estudiante es capaz de comprender información que aparece en el texto.

Debido a los resultados encontrados en los dos primeros estudios de esta tesis, diseñamos un par de experimentos en los que empleamos preguntas para evaluar el nivel de aprendizaje de conocimientos tras leer/estudiar un texto expositivo del área de

ciencias. Para ello empleamos preguntas que evaluaban no sólo la mera comprensión de la información textual, sino también la capacidad para aplicar esa información en nuevas situaciones. Es decir, ponían a prueba la habilidad del estudiante para descontextualizar la información aprendida y poder transferirla a situaciones relativamente diferentes. Imagínese que un estudiante lee en el texto el experimento de Torricelli con el que demostró que el aire ejercía presión y se le presenta la siguiente pregunta: *¿Qué quería demostrar Torricelli con su experimento?* En este caso, esta pregunta va dirigida a comprender información textual que informa sobre la presión que ejercen los gases de la atmósfera. Sin embargo, imagine que al estudiante se le plantea la siguiente pregunta: *‘Si intentas replicar el experimento de Torricelli en lo alto de una montaña de ocho mil metros, en vez de al nivel del mar, donde se realizó el original ¿Qué crees que sucederá?’*. En este caso, la pregunta no sólo requiere que el estudiante comprenda el experimento de Toricelli, sino que además razone sobre una situación hipotética que va más allá de los contenidos del texto. Así, el estudiante debe comprender que a mayor altura, existen menos capas de aire por encima, por lo que la presión ejercida por las partículas de la atmósfera será menor, dando lugar a que salga más mercurio del tubo a la cubeta, quedando por debajo de los 760mm de altura. Este tipo de preguntas evalúan el modelo de la situación construido por el estudiante, ya que le obligan a descontextualizar los conocimientos aprendidos en el texto para aplicarlos en situaciones novedosas.

En definitiva, las preguntas nos permiten conocer hasta qué punto el estudiante es capaz de comprender de forma profunda la información del texto, es decir, si es capaz de comprender y establecer relaciones entre ideas, e incluso transferir esas ideas a nuevas situaciones. Tanto las preguntas que evalúan el nivel de comprensión como las que evalúan el aprendizaje requieren de habilidades cognitivas específicas que forman parte de las características individuales de cada estudiante (e.g., competencia lectora o conocimientos previos). Estas variables intrapersonales desempeñan un papel esencial en situaciones en que se evalúa el nivel de comprensión o aprendizaje. En los dos primeros estudios tuvimos en cuenta el nivel de competencia lectora previo, medido como la habilidad para responder preguntas de comprensión de un texto disponible. En los estudios tres y cuatro evaluamos el nivel de conocimiento previo en ciencias, debido al papel que juega el conocimiento previo en la adquisición de nuevos conocimientos. En la siguiente sección analizamos el papel de ambos factores en la comprensión y aprendizaje a partir de textos.

1.2. El papel de las diferencias individuales al responder preguntas

La comprensión de textos expositivos y el aprendizaje de sus contenidos están influidos por diversos factores, entre los que destacamos las características individuales de los estudiantes. A lo largo de los últimos años se han realizado una serie de estudios para examinar el efecto de diferentes variables individuales en el proceso de responder preguntas de comprensión y aprendizaje, entre las que destacan las habilidades de comprensión y el nivel de conocimiento previo de los estudiantes. Es importante señalar que las habilidades de comprensión abarcan una amplia variedad de habilidades (e.g., acceso al léxico, comprensión de oraciones, o realización de inferencias). No obstante, a lo largo de este trabajo entenderemos habilidades de comprensión como la habilidad cognitiva general para responder preguntas de comprensión de un texto disponible, ya que ésta es la principal destreza que deben aplicar los estudiantes en las tareas administradas. Ésta es la medida empleada en los dos primeros estudios. El nivel de conocimiento previo, por su parte, hace referencia a la cantidad y la calidad de las ideas o relaciones entre ideas que el estudiante tiene sobre un dominio de conocimiento. Ésta es la medida que tomamos como referencia en los dos últimos estudios, en los que medimos el aprendizaje de conocimientos conceptuales en el área de física.

La relación entre las habilidades de comprensión y el procesamiento de la tarea y el texto al responder preguntas (i.e., creación del modelo de tarea y búsqueda de información relevante) ha sido ampliamente estudiada. Por lo que respecta al proceso de formación del modelo de tarea, el estudiante ha de procesar la información contenida en la pregunta. En ese momento, algunos estudiantes pueden experimentar dificultades para comprender la tarea de forma precisa debido, en parte, a bajas habilidades de comprensión. Vidal-Abarca et al. (2010) realizaron un experimento para examinar si los estudiantes con altas habilidades de comprensión formaban mejores modelos de tarea que los bajos en habilidades de comprensión. Para ello utilizaron el paradigma de detección de errores (e.g., Hacker, 1998; Otero, 2002) con un grupo de estudiantes de secundaria. Se presentaron una serie de preguntas con una inconsistencia interna (i.e., contenían una contradicción en su enunciado) o sin inconsistencia interna. Los estudiantes fueron informados que algunas preguntas contenían información inconsistente. En caso de identificar una pregunta con inconsistencia interna, los participantes debían contestar una versión alternativa de la pregunta. Los estudiantes

con bajas habilidades de comprensión fueron menos precisos al monitorizar su nivel de comprensión, es decir, detectaron menos preguntas inconsistentes. Según la explicación proporcionada por los investigadores, esto pudo indicar problemas para integrar la información de la pregunta (Hannon y Daneman, 2004), un proceso clave en la formación de modelos de tarea. En la misma línea, Martínez, Vidal-Abarca, Gil y Gilabert (2009) demostraron que los estudiantes con altas habilidades de comprensión relevaron las preguntas fáciles en pocas ocasiones, mientras que incrementaron las relecturas en las preguntas difíciles. Aunque los estudiantes con bajas habilidades de comprensión realizaron en promedio la misma cantidad de relecturas que los estudiantes con altas habilidades, no fueron capaces de ajustar las relecturas en función de la dificultad de las preguntas. Estas diferencias a la hora de releer la información de la pregunta en función de su dificultad sugieren que los estudiantes con buenas habilidades de comprensión son capaces de decidir cuándo es necesario actualizar el modelo de tarea. En el caso de los estudiantes con bajas habilidades de comprensión, la calidad de su modelo de tarea era tan pobre que no les permitía diferenciar cuando era necesario actualizarlo. Esto puede deberse a que el estudiante puede malinterpretar el contenido de la tarea, cometer errores al monitorizar la búsqueda de información relevante, o malinterpretar la información textual para responder. Por otra parte, Cerdán, Gilabert, y Vidal-Abarca (2013) pidieron a estudiantes con altas y bajas habilidades de comprensión que auto-explicaran la mitad de las preguntas al objeto de facilitar la comprensión profunda de su contenido. Los resultados indicaron que las auto-explicaciones de los estudiantes con bajas habilidades de comprensión incluían más ideas incompletas y erróneas, así como menos inferencias y paráfrasis que las auto-explicaciones de los estudiantes con altas habilidades de comprensión.

En cuanto a la búsqueda de información relevante en el texto, también existen evidencias que ponen de manifiesto el rol que juegan las habilidades de comprensión. Por ejemplo, Martínez et al. (2009) encontraron que los estudiantes más competentes deciden consultar el texto más a menudo cuando responden preguntas difíciles, mientras que los bajos en comprensión realizan menos consultas al texto, especialmente cuando responden preguntas difíciles. Por su parte, Cataldo y Oakhill (2000) encontraron que los estudiantes con altas habilidades de comprensión eran más estratégicos que aquellos con bajas habilidades de comprensión cuando se les pedía que buscaran información relevante para responder preguntas. Mientras que los estudiantes con altas habilidades dirigían su búsqueda a las secciones del texto donde localizar la información relevante,

los bajos en comprensión iniciaban la búsqueda desde el principio del texto. En un estudio realizado por Cerdán y Vidal-Abarca (2008), los resultados indicaron que los estudiantes con altas habilidades de comprensión ajustan la velocidad de relectura de información textual en función de su relevancia, reduciendo la velocidad de lectura cuando localizaban información relevante para la pregunta. En un estudio posterior, Cerdán et al. (2011) encontraron que los estudiantes con bajas habilidades de comprensión procesaban información no relevante durante más tiempo, aunque no encontraron diferencias en el procesamiento de información relevante.

Igualmente, los estudiantes con altas habilidades de comprensión tienden a monitorizar el uso de la información relevante de forma más precisa, ya que responden las preguntas inmediatamente después de leer un segmento de información relevante más a menudo que los estudiantes con bajas habilidades de comprensión (Gil et al., 2015; Vidal-Abarca et al., 2010, Experimento 1). Un mecanismo que puede explicar este resultado se refiere a la superposición de palabras para conectar ideas (Graesser, McNamara, Louwerse y Cai, 2004; Kintsch, 1998). Cerdán et al. (2011) encontraron que los estudiantes con bajas habilidades de comprensión guían su búsqueda a partir de señales superficiales (i.e., uso de palabras de la pregunta que también aparecen en el texto, *word-matching strategy*). Aunque los estudiantes con altas habilidades también tienden a utilizar esta estrategia inicialmente, son capaces de guiarse por contenidos semánticos y buscar información realmente relevante para la pregunta cuando la estrategia superficial falla. Ramos y Vidal-Abarca (2013) recogieron protocolos verbales de un grupo de estudiantes de secundaria para evaluar el proceso de responder preguntas de un texto. Los resultados mostraron que las verbalizaciones sobre información textual relevante de los estudiantes con altas habilidades de comprensión eran más precisas y completas que las de los estudiantes con bajas habilidades. Además, los estudiantes con altas habilidades produjeron más inferencias para conectar las ideas del texto con su conocimiento previo durante el proceso de búsqueda de información relevante.

En definitiva, estas diferencias se relacionan con procesos de comprensión que implican formar e integrar proposiciones y hacer inferencias, y que pueden jugar un papel esencial a la hora de construir el modelo de tarea, la búsqueda de información relevante, y la interacción entre la pregunta y el texto. Estas diferencias sugieren que las habilidades de comprensión juegan un papel importante al responder preguntas de

comprensión de textos expositivos, por lo que medir esta habilidad es esencial en el contexto de tarea que planteamos en los dos primeros estudios.

Por otra parte encontramos el conocimiento previo, una de las principales variables implicadas en la comprensión de textos expositivos (Afflerbach, 1986; Chi, Feltovich y Glaser, 1981; Kendeou y O'Brien, 2015, 2018; McNamara y Kintsch, 1996) y en el aprendizaje a partir de textos de ciencias (Bohn-Gettler y Kendeou, 2014; Boscolo y Mason, 2003; Cromley, Snyder-Hogan y Luciw-Dubas, 2010; Mason, Tornatora y Pluchino, 2013; Ozuru, Dempsey y McNamara, 2009; Tarchi, 2010). El conocimiento previo hace referencia al conocimiento almacenado en la memoria a largo plazo en redes de conocimiento más o menos elaboradas que el estudiante es capaz de activar de forma relativamente automática en su memoria de trabajo para poder conectarlo con nuevas ideas. El grado en que los estudiantes son capaces de conectar la información del texto con su conocimiento previo influye en el nivel de comprensión y aprendizaje de nuevos conocimientos (Kendeou et al., 2004). Investigaciones previas sugieren que la cantidad y calidad del conocimiento previo influyen en la comprensión de textos de ciencias (e.g., Boscolo y Mason, 2003; Kendeou et al., 2004; Kendeou y van den Broek, 2005, 2007; Ozuru et al., 2009; Tarchi, 2010; van den Broek, 2010), y que estos efectos pueden variar en función del tipo de texto administrado (textos refutativos vs. no refutativos) (Kendeou y van den Broek, 2007) y su nivel de coherencia (alta vs. baja coherencia) (McNamara, 2001; McNamara y Kintsch, 1996; McNamara, Kintsch, Songer y Kintsch, 1996). En una situación de aprendizaje con textos de ciencias como la que planteamos en los estudios 3 y 4 de este trabajo, el procesamiento del texto requiere que el estudiante aplique tanto procesos de comprensión como aprendizaje, ya que la nueva información ha de ser integrada en sus estructuras de conocimiento previo (van den Broek, 2010). Así, las relaciones entre las ideas del texto con el conocimiento previo activado en la memoria de trabajo del estudiante le permitirán construir una representación mental más o menos elaborada.

Mientras que los estudiantes con altos niveles de conocimiento previo son capaces de comprender de forma más precisa los textos de ciencias, probablemente porque son capaces de integrar la información de forma precisa (Mason et al., 2013), los estudiantes que disponen de conocimiento previo escaso o incluso erróneo (i.e., *misconceptions*) tienen dificultades para comprender sus contenidos, ya que su representación mental previa interfiere en la adquisición de nuevos conocimientos (e.g., Diakidoy, 1999; Diakidoy y Kendeou, 2001), e incluso puede facilitar la conexión de

ideas incorrectas (van den Broek, 2010). En el aprendizaje de conocimientos de ciencias, los lectores con bajos niveles de conocimiento previo (i.e., ideas imprecisas o erróneas) tienden a realizar inferencias incorrectas, lo que resulta en peores representaciones mentales en su memoria (Kendeou y van den Broek, 2005, 2007; van den Broek, 2010). De hecho, es bastante común que los estudiantes tengan ideas parcial o totalmente incorrectas, basadas en creencias y no en evidencia científica. Por ejemplo, los estudiantes suelen creer que, al dejar caer dos objetos de igual tamaño pero de diferente peso, el más pesado caerá antes. Esta creencia se mantiene incluso después de haber estudiado los textos en que se explican los principios de la mecánica. Los estudios presentados en este trabajo se centran en el dominio de conocimiento de física, concretamente sobre 'la presión atmosférica y el fenómeno de los vientos'. Para comprender correctamente estos conocimientos, un estudiante que dispone de mayor conocimiento de ciencias (e.g., es capaz de reconocer que la temperatura es una propiedad de la materia, o que las partículas de los gases están en constante movimiento), será capaz de construir una representación mental más integrada tras leer un texto sobre la presión atmosférica y el proceso de formación de los vientos.

Este conjunto de resultados sugiere que el conocimiento previo juega un papel esencial en el procesamiento y aprendizaje de conocimientos declarativos en el área de ciencias. Asimismo, abre el camino al desarrollo de herramientas educativas que faciliten la comprensión de los contenidos presentados en los textos expositivos (Kendeou et al., 2004). Por ejemplo, una forma de modificar conocimientos erróneos de ciencias puede ser mediante la retroalimentación formativa que incluya explicaciones sobre los fenómenos científicos evaluados mediante preguntas. Esta información puede determinar el aprendizaje de estudiantes con diferente nivel de conocimiento previo, como veremos en el apartado 2.4.

2. RETROALIMENTACIÓN FORMATIVA PARA APRENDER CONOCIMIENTO CONCEPTUAL

La retroalimentación formativa proporciona al aprendiz información sobre su rendimiento actual con respecto a un estándar para modificar sus pensamientos o comportamientos y así mejorar su aprendizaje (e.g., Narciss, 2008; Shute, 2008). Por tanto, la retroalimentación es una consecuencia derivada de su ejecución en la tarea. La retroalimentación formativa es una de las herramientas instruccionales más efectivas para mejorar el aprendizaje (Hattie y Gan, 2011; Hattie y Timperley, 2007; Shute, 2008). Proporcionar retroalimentación es una de las intervenciones psicológicas más frecuentes para mejorar la comprensión de los materiales y así fomentar el aprendizaje de los alumnos. Como se ha expuesto anteriormente, los estudiantes completan tareas académicas para aprender conocimientos de prácticamente cualquier asignatura. Tras completar estas tareas, los estudiantes suelen recibir retroalimentación de sus profesores. En el caso de contestar preguntas de un texto, el alumno ha de activar los conocimientos previos necesarios, integrar los nuevos conocimientos presentados en los materiales de aprendizaje, y revisar aquellos conocimientos erróneos o parcialmente erróneos que haya aprendido. Para ello, cada estudiante pone en marcha estrategias para resolver la tarea asignada (e.g., buscar información en los textos, discriminar la información relevante de la menos relevante, o buscar retroalimentación). Sin embargo, los estudiantes pueden carecer de suficiente conocimiento previo o codificar nuevos conocimientos de forma incompleta o incluso errónea durante la realización de las tareas. En esos casos, se espera que la retroalimentación formativa desempeñe un papel esencial, ya que se trata de una herramienta instruccional potencialmente efectiva en

gran variedad de contextos de aprendizaje (Hattie, 2009), entre los que encontramos los entornos digitales (Hattie y Gan, 2011; Hattie y Timperley, 2007; Narciss, 2013).

Uno de los primeros intentos por proporcionar retroalimentación inmediata en una máquina lo encontramos en Pressey (1926, 1950). Pressey desarrolló una máquina de enseñanza (*teaching machine*) capaz de puntuar automáticamente las respuestas de los aprendices al contestar preguntas de alternativa múltiple. Cuando el alumno proporcionaba respuestas incorrectas, la máquina mantenía la pregunta hasta que se seleccionaba la respuesta correcta (procedimiento *Try again* or *Answer-until-correct*). Por tanto, este sistema ofrecía una retroalimentación simple de verificación de respuestas junto con un procedimiento de varios intentos de respuesta. Desde una aproximación puramente conductista, Skinner (1958) desarrolló otra máquina de enseñanza capaz de proporcionar retroalimentación de forma inmediata basada en los principios del reforzamiento de conductas (i.e., los estudiantes aprendían a responder correctamente más rápido al recibir retroalimentación positiva). La máquina desarrollada por Skinner presentaba pequeños segmentos de información impresos (e.g., textos breves o ecuaciones) en una ventana, con un espacio en blanco que el alumno debía completar. Para ello, el estudiante escribía su respuesta (e.g., una palabra o un símbolo) en otra ventana en la que se presentaba un trozo de papel en blanco. Tras responder, el estudiante podía activar un mecanismo de la máquina que le informaba acerca de si su respuesta era correcta o incorrecta. Además, esta máquina permitía que el estudiante autorregulase su aprendizaje y podía ser reprogramada para ajustarse a sus necesidades en función de las ejecuciones previas.

La concepción conductista propuesta por Skinner del funcionamiento de la retroalimentación como el fortalecimiento de respuestas mediante la emisión de retroalimentación positiva al responder correctamente evolucionó con las perspectivas cognitivistas de las décadas de los 70 y 80. Esta nueva aproximación del funcionamiento de la retroalimentación se relacionó con el papel que juega esta información en la corrección de errores, más allá de la adquisición/fortalecimiento de respuestas correctas propia de la aproximación conductista (Bangert-Drowns, Kulik, Kulik y Morgan, 1991; Kulhavy, 1977; Kulhavy y Stock, 1989). Así, parece necesario que la retroalimentación sea procesada activamente para corregir errores e integrar la nueva información en la representación mental del aprendiz (e.g., Mory, 2004). Aunque las máquinas de enseñanza diseñadas en la primera mitad del siglo XX eran sencillos dispositivos mecánicos capaces de presentar preguntas y corregir respuestas

automáticamente, estos sistemas supusieron una revolución del proceso instruccional de enseñanza-aprendizaje y sentaron precedentes para los actuales entornos digitales de aprendizaje con los que podemos estudiar los procesos y estrategias que aplican los estudiantes cuando reciben retroalimentación.

El rápido desarrollo de aparatos tecnológicos y el incremento en el acceso a las nuevas tecnologías en las últimas décadas ha abierto un abanico de posibilidades en cuanto a la aplicación de tareas de aprendizaje y la entrega de retroalimentación formativa adaptada al aprendiz (e.g., Azevedo y Bernard, 1995; Mason y Bruning, 2001; Mory, 2004). Los entornos digitales de aprendizaje más avanzados (e.g., Intelligent Tutoring Systems, e-textbooks, o Massive Online Open Courses) son capaces de proporcionar retroalimentación más precisa y elaborada sobre la ejecución de los estudiantes que la proporcionada por las primeras máquinas de enseñanza e incluso por los humanos. Estos sistemas digitales pueden proporcionar pistas, explicaciones o soluciones que faciliten la comprensión de los materiales evaluados, más allá de la mera retroalimentación que únicamente informa acerca de si las respuestas son correctas o de cuál es la respuesta correcta (Tärning, 2018). Esto se debe a su capacidad para recoger y evaluar diferentes ejecuciones automáticamente con gran nivel de precisión. De este modo, estos sistemas son potencialmente eficaces para incrementar el aprendizaje de los estudiantes por medio de la entrega de retroalimentación formativa adaptada inmediatamente después de resolver una tarea (Kluger y DeNisi, 1996). Además, son capaces de recoger información conductual sobre el uso de la retroalimentación proporcionada (e.g., *¿cuánto tiempo dedican los estudiantes a procesar la retroalimentación?, ¿buscan retroalimentación adicional voluntariamente?*). Esta información permite comprender cuándo y cómo deciden emplear los estudiantes la retroalimentación al resolver tareas en entornos digitales de aprendizaje, aspecto que nos permite establecer líneas sobre la forma en que los dispositivos tecnológicos deben proporcionar retroalimentación.

A pesar de la enorme cantidad de investigación realizada sobre la efectividad de la retroalimentación a lo largo de las últimas décadas, aún se desconoce cómo se procesa la retroalimentación en el contexto de las tareas asignadas (Van der Kleij, Eggen, Timmers y Veldkamp, 2012). Las teorías recientes sobre la retroalimentación prestan poca atención a cómo los estudiantes se involucran y procesan esa información, así como el impacto que ejerce su procesamiento en la comprensión y el aprendizaje en tareas de responder preguntas a partir de textos. Los nuevos modelos de

retroalimentación deben abordar los procesos cognitivos y metacognitivos involucrados en el procesamiento de la retroalimentación en estos contextos de aprendizaje. A lo largo de este trabajo presentamos un conjunto de estudios que examinan el procesamiento de la retroalimentación formativa en un entorno digital de aprendizaje en el que los estudiantes han de contestar una serie de preguntas sobre el contenido de unos textos.

2.1. Tipos y funciones de la retroalimentación formativa

La retroalimentación formativa tiene como objetivo mejorar el conocimiento y/o las habilidades de los estudiantes, proporcionándoles información sobre ciertos aspectos de su ejecución (Hattie y Timperley, 2007). Sin embargo, la retroalimentación puede contener información muy diversa. A lo largo de los últimos años, ha habido varios intentos por crear taxonomías que organicen y clasifiquen la retroalimentación en función de varios criterios. Por ejemplo, Timmers y Veldkamp (2011) señalan tres propiedades de la retroalimentación que pueden influir en su efectividad, como son el nivel de la retroalimentación, el momento de entrega o *timing*, y el tipo de retroalimentación. Así, en función del nivel al que se dirige el contenido del mensaje, estos pueden contener información relacionada con la tarea, el proceso, la autorregulación, o uno mismo (*the task, process, self-regulation, and self levels*) (Hattie y Timperley, 2007). También se ha clasificado la retroalimentación en función del momento de entrega o *timing*, pudiendo proporcionarse los mensajes inmediatamente después de responder cada pregunta, tras responder un conjunto de preguntas, al finalizar la tarea, o tras un periodo de demora (Shute, 2008). Finalmente, la taxonomía que más nos interesa, y la que nos servirá en adelante, es la que clasifica la retroalimentación según el tipo de información de los mensajes, es decir, según su contenido (para una revisión más extensa, véase Narciss, 2013 o Shute, 2008).

En términos generales, la bibliografía suele diferenciar tres grandes tipos de retroalimentación según su contenido: (a) Verificación (Knowledge of Response o KR en inglés), (b) Respuesta Correcta (Knowledge of Correct Response o KCR en inglés), y (c) Retroalimentación Elaborada (Elaborated Feedback o EF en inglés). En adelante, emplearemos las siglas KR, KCR y EF para referirnos a estas categorías. Los mensajes KR tienen un carácter correctivo simple, incluyendo únicamente información sobre si las respuestas del alumno son correctas o incorrectas (e.g., Correcto/Incorrecto;

Verdadero/Falso; señales de verificación). Los mensajes KCR también son correctivos, aunque estos incluyen la solución correcta a una tarea específica (e.g., La respuesta/solución correcta es X; se presenta la respuesta correcta en verde y las incorrectas en rojo). Los mensajes EF, por su parte, contienen información variada sobre la precisión de las respuestas de los estudiantes como, por ejemplo, explicaciones o pistas más allá de la retroalimentación correctiva KR o KCR. Habitualmente se afirma que la retroalimentación no debe limitarse a informar únicamente del acierto/error o incluir la respuesta correcta, sino que es más útil proporcionar EF, especialmente cuando la tarea requiere procesos cognitivos de alto nivel (e.g., Butler, Godbole y Marsh, 2013; Hattie y Timperley, 2007; Kluger y DeNisi, 1996; Mory, 2004; Narciss, 2004; Shute, 2008).

Las taxonomías propuestas para la clasificación de la EF abarcan una amplia variedad de mensajes, como explicaciones, ejemplos resueltos, sugerencias o pistas, generalmente proporcionados junto con retroalimentación correctiva KR o KCR. La Tabla 1 establece una relación entre los tipos de retroalimentación definidos en las clasificaciones de Shute (2008) y Narciss (2013) según el contenido de sus mensajes.

Tabla 1. *Clasificación de los tipos de retroalimentación formativa*

Clasificación General	Shute (2008)	Narciss (2013)
KR (Verificación)	Verification (<i>Error flagging*</i> and <i>Try again**</i>)	Knowledge of performance
		Knowledge of result
KCR (Respuesta Correcta)	Correct Response	Knowledge of the correct response
EF (Retr. Elaborada)		Knowledge about task constraints
	Attribute isolation	Knowledge about concepts
	Topic contingent	
	Response contingent	
	Bugs/misconceptions	Knowledge about mistakes
	Hints/cues/prompts	
		Knowledge about meta-cognition

*Si bien Shute contempla *Error flagging* como un tipo de retroalimentación específico, consideramos que esta información entra dentro del tipo KR tal y como sugiere Narciss.

**Aunque Shute contempla *Try again* como otro tipo de retroalimentación específico, consideramos que esta información entra dentro de la categoría KR.

A lo largo del presente trabajo nos guiaremos por la clasificación general anteriormente expuesta. En general, la retroalimentación puede confirmar o modificar los conocimientos o habilidades de los estudiantes tras resolver una tarea (Mory, 2004), siendo su función principal la de promover la adquisición de conocimientos, la corrección de errores conceptuales, o la mejora de estrategias metacognitivas poco efectivas. Así, los estudiantes monitorizan activamente su rendimiento mediante la búsqueda de correspondencia entre su nivel de comprensión y el nivel proporcionado en la retroalimentación (Butler y Winne, 1995). Según estos autores, la retroalimentación tiene cinco posibles funciones según la ejecución de los aprendices: "*confirmar, sobrescribir, agregar, ajustar o reestructurar los conocimientos y creencias existentes*".

Más específicamente, podemos decir que cada tipo de retroalimentación tiene una serie de funciones propias. La retroalimentación KR permite contrastar las expectativas de rendimiento de los estudiantes, juzgar su nivel de comprensión y ser conscientes de los conocimientos erróneos previamente aprendidos. La retroalimentación KCR, además de verificar las expectativas de rendimiento de los estudiantes, les permite modificar sus conocimientos (parcialmente) erróneos por medio de la codificación del modelo de respuesta correcto. Mientras que la retroalimentación KR tiene sus orígenes en el conductismo, la retroalimentación KCR los tiene en el cognitivismo (Van der Kleij et al., 2015), ya que su función no es reforzar el recuerdo de hechos, sino la corrección de respuestas incorrectas. Sin embargo, ambos tipos de retroalimentación correctiva carecen de instrucciones que faciliten el aprendizaje. Así surge la EF, cuya función no sólo consiste en corregir errores, sino también en mejorar el rendimiento de los estudiantes mediante la comunicación de información adicional útil para la tarea. Mientras que los mensajes de retroalimentación correctiva KR y KCR requieren pocos recursos cognitivos para ser procesados, los mensajes EF requieren un procesamiento activo que permita a los estudiantes analizar sus errores a la luz de la información recibida.

Según la función correctiva de la retroalimentación, ésta se proporciona principalmente para corregir errores después de proporcionar respuestas incorrectas a tareas específicas. Esta perspectiva de la utilidad de la retroalimentación se basa en la afirmación de que las lagunas en el conocimiento deben reducirse para lograr el aprendizaje (Mory, 2004). A pesar de la veracidad de la afirmación anterior y de la supuesta ineficacia de centrarse en las respuestas correctas proporcionadas por los aprendices, consideramos que la retroalimentación dirigida a promover el aprendizaje

debe ser accesible incluso cuando los estudiantes responden correctamente, ya que pueden necesitar información correctiva o adicional para lograr una comprensión completa o correcta de los conocimientos relevantes para la tarea. En algunas ocasiones, los estudiantes pueden proporcionar respuestas correctas por mera probabilidad, lo que puede incluir imprecisiones al procesar la información, o incluso inferencias incorrectas. En otras ocasiones, los estudiantes pueden implementar estrategias de bajo nivel, como descartar respuestas incorrectas en preguntas tipo test porque no comprenden la pregunta o son incapaces de localizar información relevante para la tarea. En estas circunstancias, los estudiantes pueden no estar seguros de sus respuestas incluso cuando proporcionan la respuesta correcta. Evidencia previa sugiere que la retroalimentación puede ser efectiva para aumentar la retención de respuestas correctas con bajos niveles de confianza, lo que significa que la retroalimentación es capaz no sólo de corregir errores conceptuales sino también de corregir imprecisiones metacognitivas para preguntas respondidas correctamente (Butler, Karpicke y Roediger, 2008). Por lo tanto, el modelo de respuesta del estudiante tiene que ser desafiado o verificado con el fin de mejorar la corrección del conocimiento erróneo o los juicios metacognitivos inexactos. Incluso en las situaciones en que se proporciona retroalimentación adecuada en el momento preciso, cabe señalar que su efectividad no es inmediata. Para que resulte efectiva, los estudiantes deben participar activamente en el procesamiento de su contenido al objeto de que la información pueda ser integrada en su representación mental de la temática evaluada (i.e., los alumnos revisan y actualizan sus conocimientos o estrategias en función de la retroalimentación recibida), lo que significa que deben implementar procesos cognitivos y metacognitivos (Bangert-Drowns et al., 1991; Carless y Boud, 2018).

Aunque no existe consenso en la comunidad científica sobre los tipos de información que deben incluirse en la retroalimentación, los autores coinciden en afirmar que la retroalimentación formativa debe integrar información correctiva y elaborada (Kulhavy y Stock, 1989; Mason y Brunning, 2001), estableciendo relaciones entre la ejecución del estudiante y el estándar diseñado para la tarea en cuestión (e.g., Kluger y DeNisi, 1996; Narciss y Huth, 2004). De hecho, parece que la retroalimentación que contiene información sobre la tarea y cómo realizarla de manera más efectiva juega un papel esencial en el aprendizaje (Hattie y Timperley, 2007). Ahora bien, ¿qué tipo de retroalimentación formativa es más eficaz para facilitar el aprendizaje de conocimientos conceptuales? En la siguiente sección revisamos la

efectividad de la retroalimentación formativa en tareas de aprendizaje, prestando especial atención al aprendizaje que requiere procesamiento cognitivo de alto nivel.

2.2. Efectividad de la retroalimentación formativa

Una de las cuestiones más ampliamente estudiadas en el área de la retroalimentación hace referencia a su efectividad en el aprendizaje de conocimientos y competencias. Estudios teóricos y meta-análisis clásicos consideran que la retroalimentación puede producir efectos positivos en el aprendizaje, aunque sus efectos suelen ser inconsistentes y variables (e.g., Azevedo y Bernard, 1995; Jaehnig y Miller, 2007; Kluger y DeNisi, 1996; Kulhavy, White, Topp, Chan y Adams, 1985; Shute, 2008; Van der Kleij, Timmers y Eggen, 2011). En el caso de las tareas de responder preguntas de un texto, la retroalimentación formativa puede tener un impacto en la comprensión y el aprendizaje de conceptos e ideas.

Aunque no existe un consenso claro sobre el papel que juega la retroalimentación en el aprendizaje (Narciss y Huth, 2004), resultados de un reciente meta-análisis sobre la efectividad de la retroalimentación en entornos digitales de aprendizaje ponen de manifiesto que la EF es la más efectiva (Van der Kleij et al., 2015), especialmente cuando se trata de tareas de aprendizaje de orden superior, ya que su contenido permite a los estudiantes acortar la distancia entre el nivel de rendimiento alcanzado y el deseado. Van der Kleij et al. (2015) encontraron que el tamaño del efecto de la retroalimentación KR, KCR, y EF es .05, .32 y .49, respectivamente. Por tanto, la retroalimentación KCR produce resultados relativamente positivos, mientras que la retroalimentación KR no produce efectos en el rendimiento, ya que no fomenta la construcción de significado (Hattie y Timperley, 2007). Así, es posible que la retroalimentación correctiva no produzca resultados tan positivos en el aprendizaje porque su información no facilita la integración del nuevo conocimiento en las representaciones mentales previamente construidas. Proporcionar EF es lo más aconsejable, aunque para ello parece necesario hacer uso de entornos digitales que puedan proporcionar esta información de forma individual e inmediata. Es importante recordar que el concepto EF engloba gran variedad de mensajes (e.g., explicaciones, pistas, ejemplos) y que actualmente se desconoce qué subtipo es más efectivo (Golke, Dörfler y Artelt, 2015). El presente trabajo no pretende responder esta cuestión, sino sólo examinar hasta qué punto están dispuestos los estudiantes de secundaria a procesar

mensajes EF al responder preguntas de un texto en un entorno digital de aprendizaje. Asimismo, el segundo estudio explora los procesos (meta)cognitivos que intervienen en su procesamiento.

Habitualmente, las evaluaciones con retroalimentación diseñadas para medir el aprendizaje potencial de los estudiantes utilizan dos diseños: test-entrenamiento-test (*test-train-test design*) y entrenamiento-intra-test (*train-within-test design*) (Dillon, 1997, en Dörfler, Golke y Artelt, 2017). En las evaluaciones test-entrenamiento-test, el efecto de la retroalimentación se mide en el rendimiento obtenido en un test aplicado al finalizar la fase de entrenamiento. Sin embargo, en las evaluaciones entrenamiento-intra-test, el efecto de la retroalimentación se mide en el rendimiento obtenido al completar el propio entrenamiento. Aunque ambas evaluaciones son ampliamente utilizadas en el contexto educativo, las investigaciones que evalúan la efectividad de la retroalimentación tienden a utilizar el diseño test-entrenamiento-test. En el presente trabajo, el primer estudio evalúa el impacto de la retroalimentación en el rendimiento de los participantes al completar el propio entrenamiento. Sin embargo, los estudios tres y cuatro evalúan su efecto en el rendimiento obtenido tanto en la fase de entrenamiento como en una tarea final administrada tras un periodo de demora.

Basándonos en la literatura existente, la efectividad de la retroalimentación ha sido ampliamente estudiada en las últimas décadas en una amplia variedad de escenarios y tareas de aprendizaje. El aprendizaje de conocimientos y competencias puede ser más o menos complejo en función de las características y requisitos de las tareas asignadas (Narciss, 2013). Una distinción general podría realizarse en función del nivel de procesamiento que requieren los materiales de aprendizaje utilizados. Mientras que algunos estudios han evaluado la efectividad de la retroalimentación en el aprendizaje de materiales sencillos como vocabulario, definiciones de conceptos o hechos generales haciendo uso de "evaluaciones con retroalimentación" (*testing with feedback*) (e.g., Butler et al., 2008; Lipko-Speed, Dunlosky y Rawson, 2014), otros han examinado su efectividad utilizando tareas más complejas y desafiantes como, por ejemplo, responder preguntas de comprensión de textos (e.g., Butler et al. 2013; Golke et al., 2015; Llorens, Vidal-Abarca y Cerdán, 2016; Murphy, 2007) o aprender conocimientos de algún dominio académico como ciencias o matemáticas (e.g., Attali y Van der Kleij, 2017; Corbalan, Paas y Cuypers, 2010; Lee, Lim y Grabowski, 2009; Maier, Wolf y Randler, 2016; Moreno, 2004; Rakoczy, Harks, Klieme, Blum y Hochweber, 2013). Aunque en el presente trabajo nos centraremos en tareas que requieren procesos de orden superior

(e.g., realizar inferencias o aplicar conocimientos a situaciones nuevas), a continuación resumimos algunos trabajos relevantes para ilustrar la efectividad de varios tipos de retroalimentación en tareas sencillas y complejas administradas en entornos digitales de aprendizaje.

Investigaciones recientes han examinado los efectos de las evaluaciones con retroalimentación (*testing with feedback*) en el aprendizaje de definiciones y conceptos sencillos por parte de estudiantes jóvenes (e.g., Lipko-Speed et al., 2014; Metcalfe y Kornell, 2007; Metcalfe, Kornell y Finn, 2009; Metcalfe, Kornell y Son, 2007). En general, estos estudios han encontrado que los estudiantes que completan evaluaciones seguidas de retroalimentación aumentan su rendimiento de aprendizaje, medido como la retención a largo plazo de materiales simples, en comparación con completar las evaluaciones sin retroalimentación o sólo estudiar los materiales (Butler et al., 2008; Butler y Roediger, 2008; Dunlosky, Rawson, Marsh, Nathan y Willingham, 2013; Karpicke y Roediger, 2007). Por ejemplo, Lipko-Speed et al. (2014) realizaron dos experimentos para investigar los efectos de las evaluaciones con retroalimentación (i.e., evaluaciones seguidas de reestudio de las definiciones correctas) en el aprendizaje de conceptos de ciencias en un entorno digital de aprendizaje. Un grupo de estudiantes de quinto grado estudiaron veinte definiciones de conceptos científicos sobre la luz y el sonido en el Experimento 1 (e.g., *¿Qué es el sonido?*) y sobre geografía en el Experimento 2 (e.g., *¿Qué es el clima?*). Después de completar la sesión de estudio inicial, las definiciones fueron asignadas aleatoriamente a una de cuatro condiciones experimentales: reestudio de las definiciones, evaluación sin retroalimentación, evaluación con retroalimentación, o grupo control (sin estudio o evaluación adicional). Finalmente, los estudiantes completaron una prueba final en la que tuvieron que escribir la definición de cada concepto científico. Los resultados mostraron que las evaluaciones con retroalimentación mejoraron el rendimiento de los estudiantes en una tarea de aprendizaje de conceptos científicos.

Un resultado interesante en el campo de las evaluaciones con retroalimentación se refiere al hecho de que proveer retroalimentación es crucial para mejorar el aprendizaje de respuestas correctas, especialmente cuando se proporcionan de forma dubitativa (i.e., baja seguridad). Butler et al. (2008) pidieron a estudiantes universitarios que respondieran una serie de preguntas que evaluaban conocimientos generales. Para cada pregunta, los participantes tuvieron que seleccionar una respuesta y evaluar su grado de confianza en la misma. La mitad de las preguntas fueron seguidas por

retroalimentación KCR, mientras que la otra mitad fueron evaluadas sin retroalimentación. Posteriormente, los participantes completaron una prueba final que incluía el conjunto de preguntas respondidas previamente bajo las condiciones de retroalimentación KCR y sin retroalimentación, así como un conjunto de preguntas no evaluadas (condición control). Butler et al. (2008) encontraron que la retroalimentación KCR mejoró la retención de respuestas correctas no sólo en las preguntas respondidas incorrectamente (i.e., aprendizaje de respuestas correctas), sino también en las preguntas respondidas correctamente con baja seguridad (i.e., fortalecimiento de respuestas correctas). Este hallazgo sugiere que la utilidad de la retroalimentación va más allá de su función básica de corrección de errores. Por tanto, la retroalimentación no debe proporcionarse únicamente tras responder incorrectamente (como se hace con frecuencia en investigación), sino que también debe proporcionarse después de dar respuestas correctas.

La retroalimentación proporcionada en estos estudios (*testing with feedback*) se reduce generalmente a la revisión de las respuestas correctas, es decir, al reestudio de los materiales tras completar la evaluación. Por tanto, dado que los materiales de aprendizaje se presentan nuevamente, los alumnos tienen acceso a las definiciones correctas, es decir, retroalimentación KCR. No obstante, no reciben retroalimentación KR o EF. En tales casos, los estudiantes pueden cometer errores al monitorizar la precisión de sus definiciones con respecto al nivel deseado de rendimiento, ya que generalmente experimentan dificultades para monitorizar su aprendizaje (Lipko-Speed et al., 2009), especialmente cuando los materiales de aprendizaje son complejos (e.g., Dunlosky y Lipko-Speed, 2007; Thiede, Griffin, Wiley y Redford, 2009). Puede que los estudiantes experimenten dificultades para distinguir entre diferentes grados de precisión y, como consecuencia, no sean capaces de dedicar más recursos cognitivos a respuestas incompletas o incorrectas, especialmente cuando las tareas de aprendizaje requieren un procesamiento profundo. Como se indicó anteriormente, recibir retroalimentación KR podría disminuir el impacto de la monitorización inexacta sobre la percepción de la precisión de las respuestas (Butler et al., 2008). Además, Lipko-Speed et al. (2014) encontraron que los estudiantes reestudiaron las definiciones durante la retroalimentación de forma superficial. Esto pudo deberse al hecho de que la retroalimentación no tuvo en cuenta ninguna información correctiva o adicional para comprender mejor los conceptos estudiados. No obstante lo anterior, las evaluaciones con retroalimentación benefician la recuperación de conocimientos por parte de los

estudiantes e incluso facilitan la transferencia de conocimientos (Carpenter, 2012; Karpicke, 2012). Estudios más recientes han explorado el papel de la EF en el aprendizaje de conceptos. Por ejemplo, Finn, Thomas y Rawson (2018) encontraron que proporcionar EF que incluya ejemplos sobre los conceptos aprendidos mejora significativamente el aprendizaje.

Aunque las tareas presentadas anteriormente puedan resultar difíciles para los participantes, generalmente implican un aprendizaje asociativo o memorístico de hechos o conceptos en lugar de un aprendizaje complejo que requiera un procesamiento profundo de la información. En estos últimos escenarios, el papel de la EF adquiere un protagonismo mayor (e.g., Bangert-Drowns et al., 1991), aunque su efectividad puede ser menor debido a la dificultad de corregir e integrar el nuevo conocimiento en las representaciones mentales de los estudiantes. La efectividad de la retroalimentación también se ha estudiado en estas situaciones de aprendizaje de conocimientos y desarrollo de habilidades que generalmente involucran la comprensión de textos, el aprendizaje de materiales de ciencias o la resolución de problemas matemáticos. Estos entornos de aprendizaje requieren que los estudiantes lean para comprender, una actividad cognitiva compleja que en ocasiones requiere intervenciones de larga duración (Dörfler et al., 2017). En estas situaciones de evaluación, la mera memorización del contenido de los materiales de aprendizaje no mejoraría la comprensión profunda de la información.

Estudios recientes han examinado los efectos de la retroalimentación formativa en la comprensión de textos (e.g., Golke et al., 2015; Lee et al., 2009; Murphy, 2007; Schunk y Rice, 1991) y en el uso de estrategias de lectura (e.g., buscar información en los textos) en entornos digitales de aprendizaje (e.g., Llorens, Cerdán y Vidal-Abarca, 2014; Llorens et al., 2016; Llorens, Vidal-Abarca, Cerdán y Ávila, 2015). Estos estudios han arrojado resultados mixtos sobre la efectividad de la retroalimentación. Mientras que algunos no encontraron efectos positivos de la EF en la comprensión de los textos (Golke et al., 2015; Murphy, 2007), otros encontraron efectos positivos (Schunk y Rice, 1991), especialmente cuando la EF incluye información sobre las conductas de búsqueda de información en el texto (Llorens et al., 2016). Estos estudios demuestran que el aprendizaje a partir de textos conlleva una alta demanda de procesamiento de la información y que los estudiantes han de percibir la retroalimentación como un componente útil para aprender (Dörfler et al., 2017). Aunque la retroalimentación es un apoyo instruccional que debería ayudar a la comprensión de

los materiales (Langer, Keenan y Medosch-Schonbeck, 1985), aún se desconoce cómo su presencia afecta al procesamiento de los textos (Langer y Keenan, 2000).

Golke et al. (2015) examinaron los efectos de diferentes tipos de EF en la comprensión de textos con estudiantes de sexto grado. Tres tipos de EF fueron empleados. Estos mensajes estaban formados por retroalimentación KR junto con apoyos para realizar inferencias (i.e., una ayuda para mejorar la construcción de inferencias requeridas por la pregunta), explicaciones del error (i.e., la razón por la cual una respuesta era incorrecta), y ayudas de monitorización (i.e., un apoyo para monitorizar la comprensión del texto). La efectividad de estos tipos de EF fue comparada con una condición que sólo incluía retroalimentación KR y una condición sin retroalimentación. Los participantes leyeron varios textos expositivos y narrativos para responder preguntas de comprensión tipo test en un entorno digital. Todos los estudiantes, excepto los del grupo sin retroalimentación, recibieron los mensajes asignados después de contestar incorrectamente cada pregunta. Después de recibir esta retroalimentación, tuvieron un segundo intento sobre el que no recibieron retroalimentación. Los participantes completaron una evaluación inmediatamente después de finalizar la prueba experimental y otra varias semanas después. Ambas pruebas incluyeron diferentes textos y preguntas. Los resultados mostraron que ni el rendimiento de comprensión ni el tiempo dedicado a responder las preguntas diferían entre los tipos de retroalimentación. En otras palabras, ninguno de los tipos de EF empleados aumentó el rendimiento o influyó en el tiempo dedicado a completar la tarea en comparación con la retroalimentación KR o el grupo sin retroalimentación. Además, la retroalimentación KR tampoco afectó a la comprensión de los estudiantes. Golke et al. (2015) sugirieron que los estudiantes no se involucraron activamente en el procesamiento de la EF para mejorar su comprensión, ya que no dedicaron mucho tiempo a procesar los mensajes de retroalimentación o buscar información en el texto antes del segundo intento. En su segundo experimento, Golke et al. (2015) examinaron si la efectividad de la EF puede estar influida por la forma de presentación (ordenador vs. persona). En este caso, la EF entregada después de dar respuestas incorrectas fue proporcionada por una persona o por un ordenador. Los resultados mostraron que la EF proporcionada por personas fue más efectiva que cualquiera de los grupos de retroalimentación proporcionada por el ordenador para mejorar la comprensión de los estudiantes. El rendimiento fue similar entre los grupos de retroalimentación proporcionada por ordenador, igual que sucedió en el primer estudio. En un estudio

previo, Kluger y Adler (1993) encontraron que los estudiantes tienden a buscar retroalimentación más a menudo cuando es proporcionada por ordenadores en comparación con la proporcionada por personas.

Otro estudio relevante sobre los efectos de la retroalimentación en el aprendizaje a partir de textos es el realizado por Butler et al. (2013), quienes llevaron a cabo dos experimentos en los que se manipuló el tipo de retroalimentación para examinar si la EF consistente en proporcionar explicaciones de la respuesta correcta facilita la transferencia del aprendizaje. Un grupo de estudiantes universitarios estudiaron 10 textos y contestaron 20 preguntas sobre definiciones sin acceso a los materiales de estudio. Después de responder cada pregunta, los participantes recibieron retroalimentación según la condición asignada: KCR, EF con explicaciones, o control sin retroalimentación. Dos días después, se midió la transferencia de aprendizaje utilizando preguntas inferenciales nuevas y repetidas. Los resultados indicaron que el rendimiento en las preguntas repetidas no difirió entre la retroalimentación KCR y la EF, mientras que el rendimiento en las preguntas inferenciales nuevas aumentó en la condición EF. Butler et al. (2013) concluyeron que el contenido de la retroalimentación es el factor más importante de la eficacia de la retroalimentación, lo que sugiere que la EF basada en explicaciones es más beneficiosa que la retroalimentación KCR en la transferencia del aprendizaje.

Investigaciones recientes han demostrado que la EF que informa sobre la estrategia correcta al responder preguntas de un texto mejora tanto la comprensión de los estudiantes de secundaria como sus habilidades de lectura (e.g., estrategias de búsqueda de información, es decir, *cuándo* y *qué* información buscar) (e.g., Llorens et al., 2014; Llorens et al., 2016; Llorens et al., 2015). Las decisiones ilustradas en el modelo de procesamiento de tareas (Figura 1) pueden ser fácilmente almacenadas por sistemas informáticos para, posteriormente, ser transformadas en mensajes EF que permitan al estudiante modificar sus estrategias al contestar preguntas, así como actualizar su representación mental de los conocimientos evaluados. Llorens et al. (2014) analizaron el impacto de la retroalimentación formativa en las decisiones de búsqueda de información en el texto. Estos investigadores generaron un conjunto de mensajes EF que incluían retroalimentación KCR junto con información sobre los comportamientos de búsqueda de los estudiantes al responder cada pregunta (i.e., *cuándo* y *qué* información textual habían buscado los estudiantes). Los resultados indicaron que esta EF que incluía información sobre el proceso de búsqueda mejoraba el

rendimiento en comprensión y la búsqueda estratégica de información textual en comparación con la retroalimentación que sólo incluía KCR y la que contenía información general sobre las decisiones de búsqueda. Posteriormente, Llorens et al. (2015) demostraron que proporcionar EF indicando *qué* información textual es relevante para la pregunta mejoró el rendimiento de los estudiantes y los comportamientos de búsqueda de información textual. Es posible que la efectividad de esta EF se deba a que permite hacer comparaciones entre el nivel de rendimiento actual de los estudiantes (i.e., información que he buscado) y el nivel deseado (i.e., información que debería haber encontrado en mi búsqueda).

Basándose en resultados previos, Llorens et al. (2016) llevaron a cabo dos experimentos para examinar los efectos de la EF en las estrategias de lectura autorreguladas y el rendimiento en comprensión. Un grupo de estudiantes de secundaria respondieron una serie de preguntas tipo test con el texto disponible. Se aplicaron dos procedimientos que diferían en la tarea y en la retroalimentación proporcionada. En una condición, los estudiantes respondieron las preguntas pudiendo buscar información en el texto cuando lo consideraran oportuno. En la otra condición, sin embargo, los estudiantes fueron forzados a buscar y seleccionar la información relevante para responder cada pregunta. Los estudiantes recibieron EF que incluía retroalimentación KR y KCR sobre la precisión de sus respuestas y el procesamiento del texto (e.g., información relevante leída o seleccionada, dependiendo de la condición). Los investigadores incluyeron un grupo control en el que los participantes respondieron las preguntas con el texto disponible pero no recibieron retroalimentación. Posteriormente, los participantes completaron una tarea final sin retroalimentación. Los resultados mostraron que los estudiantes que tenían que seleccionar información relevante y recibían EF aplicaban estrategias de búsqueda (i.e., *cuándo* buscar y *qué* información buscar) más a menudo, lo que a su vez tenía un impacto positivo en su comprensión. Llorens et al. (2016) realizaron un segundo experimento para investigar los efectos específicos de la retroalimentación y la tarea de selección de información. Los estudiantes tenían que seleccionar la información que consideraban relevante para cada pregunta. Mientras que un grupo de estudiantes recibió la EF explicada anteriormente, el otro grupo no recibió retroalimentación. Los resultados indicaron que tanto la retroalimentación como la tarea de selección jugaron papeles importantes en la comprensión de los estudiantes. En estos casos, parece que proporcionar EF con la respuesta correcta o las ideas esenciales que debe contener una respuesta es un

procedimiento efectivo para mejorar el proceso de monitorización en la tarea (Dunlosky, Hartwig, Rawson y Lipko-Speed, 2010; Rawson y Dunlosky, 2007).

Respecto al aprendizaje de conocimientos conceptuales en el área de ciencias, a lo largo de los últimos años se han realizado numerosos estudios para evaluar los efectos de la retroalimentación formativa en entornos digitales de aprendizaje (e.g., Lee et al., 2009; Maier et al., 2016; Moreno, 2004; Smits, Boon, Sluijsmans y van Gog, 2008). Por ejemplo, Moreno (2004) estudió si la EF basada en explicaciones funcionaba mejor que la retroalimentación correctiva en una situación en la que los aprendices debían diseñar una planta capaz de sobrevivir en diferentes condiciones climáticas. El grupo de estudiantes que recibió EF obtuvo mejor rendimiento en una tarea de transferencia aplicada tras un periodo de demora. Según señala Moreno (2004), es posible que las explicaciones redujeran la carga cognitiva de la tarea.

Lee et al. (2009) examinaron los efectos de las estrategias de aprendizaje y la retroalimentación formativa en la comprensión de un texto sobre el corazón humano. Estudiantes universitarios estudiaron un texto sobre el corazón humano bajo una de las tres condiciones experimentales: KR, EF, o control. Los participantes en las dos primeras condiciones leyeron el texto y realizaron las mismas tareas: resaltar, tomar notas y responder preguntas. La única diferencia fue la retroalimentación proporcionada después de responder cada pregunta. Mientras que los participantes del grupo KR recibieron retroalimentación sobre la corrección de sus respuestas (*Correcto o Incorrecto*), los participantes del grupo EF recibieron retroalimentación KR junto con un mensaje metacognitivo (e.g., *¡Incorrecto! Necesitas volver y revisar la información destacada o el resumen*). A los participantes del grupo control se les presentó el texto, pero no se les administró ninguna tarea ni retroalimentación adicional. Se midió el conocimiento previo y la comprensión final posterior a la fase de aprendizaje. Además, los participantes fueron preguntados por sus estrategias de aprendizaje autorregulado. Los resultados mostraron que proporcionar EF mejoró los resultados de comprensión y las estrategias en comparación con el grupo control. Sin embargo, no se encontraron diferencias entre los grupos KR y EF. Es importante mencionar que Lee et al. (2009) tuvieron treinta y seis participantes para un diseño de 3x2, con un número pequeño de participantes en cada celda (4-9 estudiantes). Similares resultados fueron encontrados por estos mismos autores en un estudio posterior (Lee, Lim y Grabowski, 2010) en el que los estudiantes que recibieron EF con información metacognitiva junto con pistas sobre las estrategias de aprendizaje mejoraron su comprensión del texto sobre el

funcionamiento del corazón humano. En este caso, participaron 223 estudiantes universitarios.

También en el área de ciencias, Smits et al. (2008) diseñaron un estudio en el que un grupo de estudiantes de secundaria resolvió cuatro problemas sobre genética en un entorno digital de aprendizaje, pudiendo recibir retroalimentación global o EF (i.e., según su contenido) de forma inmediata o demorada (i.e., según el *timing*). Aquellos que recibieron retroalimentación global fueron informados sobre la precisión de sus respuestas junto con un mensaje general que incluía los pasos para solucionar problemas. Además de la información anterior, los estudiantes que recibieron EF también fueron informados sobre los pasos resueltos para solucionar el problema junto con explicaciones sobre la respuesta correcta. La retroalimentación fue proporcionada inmediatamente después de completar cada tarea o después de resolver dos tareas. Los participantes completaron un pretest que sirvió para dividir a la muestra en estudiantes con alto y bajo conocimiento previo. Después de realizar las tareas de aprendizaje con retroalimentación, los estudiantes completaron un posttest con nuevas preguntas en el ordenador. Los resultados mostraron que los estudiantes con bajo conocimiento previo obtuvieron resultados similares en el posttest independientemente de la retroalimentación recibida. Sin embargo, los estudiantes con alto conocimiento previo obtuvieron un mejor rendimiento cuando recibieron retroalimentación global en comparación con la EF. Smits et al. (2008) concluyeron que los estudiantes con bajo conocimiento previo tal vez no se beneficiaron de la EF proporcionada inmediatamente porque ya disponían de cierto conocimiento previo. Asimismo, es posible que los estudiantes con alto conocimiento previo se beneficiaran de la retroalimentación global debido a que su conocimiento les permitió elaborar sobre la información sin necesidad de recibir retroalimentación más específica.

Recientemente, Maier et al. (2016) realizaron un estudio para analizar la efectividad de la retroalimentación formativa en el aprendizaje de conocimientos de ciencias. Los autores incluyeron tres condiciones según el tipo de retroalimentación: EF (i.e., KR + explicación tras responder incorrectamente), KR, y control sin retroalimentación. Estudiantes de secundaria leyeron un texto de biología, contestaron una serie de preguntas y recibieron retroalimentación según el grupo asignado. Además, pidieron a los alumnos que recibieron EF que autoevaluaran el uso que habían realizado de esa información. Los resultados iniciales mostraron que no hubo diferencias entre los tres tipos de retroalimentación en el rendimiento de los estudiantes. Posteriormente,

Maier et al. (2016) decidieron dividir a los estudiantes que recibieron EF en dos subgrupos: aquellos que indicaron haber utilizado esa información y aquellos que indicaron no haberla utilizado. Los nuevos resultados mostraron que aquellos estudiantes que decían haber utilizado la EF obtuvieron mejor rendimiento que los estudiantes del grupo control y aquellos que indicaron no haberla utilizado.

Es importante mencionar que los estudios presentados anteriormente incluyeron diferentes tipos de EF, bien formada por información de tipo metacognitivo (e.g., pistas para monitorizar la comprensión), bien por información de tipo cognitivo (e.g., pistas para generar inferencias). La retroalimentación metacognitiva generalmente informa a los estudiantes sobre qué estrategias cognitivas implementar y cómo usarlas de manera apropiada (Butler y Winne, 1995), haciendo conscientes a los estudiantes de las estrategias de aprendizaje utilizadas (Lee et al., 2009). Sin embargo, la retroalimentación cognitiva proporciona información sobre el proceso cognitivo que el alumno debe implementar para comprender con éxito el conocimiento evaluado (e.g., información para generar inferencias). Un procedimiento habitual consiste en proporcionar EF con explicaciones que orienten al alumno hacia la comprensión de la respuesta correcta. Los estudios incluidos en esta tesis incluyeron ambos tipos de información.

Así, en los dos primeros estudios de esta tesis los estudiantes tenían que seleccionar la información que consideraran relevante para responder las preguntas de un texto. En estos casos, los estudiantes recibieron EF que contenía información tanto cognitiva (qué información es relevante y no relevante) como metacognitiva (pistas de monitorización para fomentar las comparaciones entre su selección y la selección correcta). Un ejemplo de esta retroalimentación podría ser el siguiente: "*Has seleccionado información no relevante. Revisa tu selección de acuerdo a la selección correcta. La próxima vez intenta seleccionar sólo la información relevante para la pregunta*". Sin embargo, en los estudios tres y cuatro, los estudiantes recibieron EF de tipo cognitivo formada por explicaciones para facilitar las inferencias que ayuden a comprender la respuesta correcta, pero sin presentarla explícitamente. Por ejemplo, si preguntamos a un estudiante acerca de la influencia de la altitud en el resultado del experimento de Torricelli sobre la presión atmosférica: "*Si intentas replicar el experimento de Torricelli en lo alto de una montaña de ocho mil metros en lugar de a nivel del mar, donde Torricelli realizó el original, ¿qué crees que sucederá?*" En este caso, los estudiantes deben comprender el experimento de Torricelli para descubrir que

en la cima de la montaña la densidad del aire es más baja que a nivel del mar y, por consiguiente, el peso de las capas de aire es menor, por lo que se necesita menos mercurio para igualar la presión del aire. Para esta pregunta, el mensaje EF empleado fue el siguiente: *"Piensa que como estás a mayor altitud, hay menos aire por encima. Por tanto, el peso del aire y el mercurio dentro del tubo se igualarán inmediatamente después de que haya salido una cantidad diferente de mercurio en lugar de cuando se realiza a nivel del mar"*. Si el estudiante está dispuesto y es capaz de procesar esta retroalimentación, entonces podrá comprender que *"Saldrá más mercurio del tubo al cubo, descendiendo más de 760 mm de altura"*. Es importante señalar que en los cuatro estudios la EF fue administrada junto con retroalimentación correctiva para examinar a qué componentes de la retroalimentación formativa prestan atención los estudiantes de secundaria, y cómo esta información influye en el uso de la EF.

En entornos digitales de aprendizaje, es importante tener en cuenta que proporcionar mensajes de retroalimentación no significa necesariamente que los estudiantes procesen detenidamente su información (Aleven, Stahl, Schworm, Fischer y Wallace, 2003). De hecho, es más probable que los estudiantes decidan no procesar la retroalimentación cuando ésta es proporcionada por un sistema informático que por un humano. Sin embargo, los entornos digitales ofrecen la posibilidad de controlar el proceso de aprendizaje más en detalle, ya que son capaces de registrar datos sobre la interacción del alumno con los materiales en tiempo real para, posteriormente, proporcionar retroalimentación formativa en función de su ejecución. Igualmente, las acciones de los estudiantes proporcionan información acerca del uso y procesamiento de la retroalimentación (e.g., buscan retroalimentación, cierran los mensajes sin procesarlos, cuánto tiempo dedican a los mensajes). Esa información puede ser empleada para diseñar entornos digitales de aprendizaje potencialmente más eficaces. En el caso de los libros de texto electrónicos (*e-textbooks*), estos pueden ser diseñados para ofrecer retroalimentación bajo demanda del estudiante. Así, los estudiantes pueden autorregular el uso de la retroalimentación en función de si consideran necesario acceder a esos mensajes o no. Recientemente, se han llevado a cabo estudios en los que los participantes pueden usar la EF voluntariamente, es decir, ellos deciden si acceden o no a esos mensajes (e.g., Fox, Klein Entink y Timmers, 2014; Timmers y Veldkamp, 2011). Por ejemplo, Timmers y Veldkamp (2011) diseñaron una evaluación formativa en un entorno digital en la que los participantes podían acceder opcionalmente a mensajes EF que incluían retroalimentación KCR y una explicación de los conceptos

evaluados después de contestar una prueba sobre competencia en el uso de información y recibir retroalimentación KR para cada respuesta. Los resultados de este estudio mostraron que los participantes acceden y procesan los mensajes EF opcionales en función del resultado de sus respuestas, es decir, acceden a más mensajes para las preguntas respondidas incorrectamente.

En conclusión, la efectividad de la retroalimentación formativa se ha evaluado en una amplia variedad de escenarios en entornos digitales. Una vertiente amplia del campo de estudio de la retroalimentación ha examinado sus efectos en el aprendizaje de conocimientos sencillos como, por ejemplo, vocabulario o definiciones. En estas situaciones de aprendizaje, parece que la retroalimentación, entendida como la presentación y reestudio de los materiales evaluados (i.e., retroalimentación KCR), tiene efectos positivos en el rendimiento de los estudiantes. Sin embargo, si nos centramos en evaluaciones de materiales más complejos, generalmente relacionados con habilidades de comprensión lectora o dominios de conocimiento como ciencias, que requieren la aplicación de procesos cognitivos superiores (e.g., realización de inferencias o aplicación de los conocimientos aprendidos a nuevas situaciones), los efectos de la retroalimentación muestran resultados altamente variables. Estos escenarios guardan una estrecha relación con el aprendizaje de conocimientos conceptuales a partir de textos en el ámbito académico. Aunque en términos generales la EF es el tipo de retroalimentación más efectiva en entornos digitales de aprendizaje, todavía existen lagunas sobre el uso y procesamiento que realizan los estudiantes cuando reciben este tipo de retroalimentación al responder preguntas de comprensión y aprendizaje sobre los contenidos de un texto. En la siguiente sección revisamos los principales modelos teóricos sobre retroalimentación y explicamos qué significa, cognitivamente y metacognitivamente, procesar la retroalimentación en el contexto de tarea explicado anteriormente.

2.3. Marco de trabajo para el procesamiento de la retroalimentación

La retroalimentación es una herramienta de enseñanza ampliamente utilizada para fomentar el aprendizaje de los estudiantes. A pesar de su potencial eficacia para mejorar el rendimiento, los investigadores y profesionales del área del aprendizaje parecen asumir que todos los estudiantes procesan la información adecuadamente una vez recibida (Corbalan, Kester y Van Merriënboer, 2009; Gordijn y Nijhof, 2002; Smits

et al., 2008; Van der Kleij et al., 2011). Investigaciones previas demuestran que no podemos asumir que los alumnos harán un uso adecuado de esta información (Sadler, 1998). Como sugiere Langer (2011), la retroalimentación debe conceptualizarse como un proceso de resolución de problemas que debe ajustarse al estudiante. Al recibir retroalimentación, los estudiantes han de evaluar la precisión de sus respuestas activamente al objeto de ajustar su representación al estándar proporcionado (e.g., Bangert-Drowns et al., 1991), por lo que la retroalimentación puede tener un significado muy diferente en función de la interpretación que haga cada estudiante (Black y Wiliam, 2009).

Esta perspectiva reconoce el papel activo de los estudiantes, quienes deben implicarse en un proceso de resolución de problemas al recibir la retroalimentación. En ese momento, los estudiantes necesitan comprender la información contenida en el mensaje y tomar decisiones sobre cómo procesar la información recibida. Así, cuando los estudiantes operan en entornos en los que han de autorregular su aprendizaje, como es el caso de las tareas de responder preguntas a partir de textos con retroalimentación formativa, los estudiantes deben monitorizar, regular y controlar su cognición, motivación y comportamiento (Pintrich y Zusho, 2002). En otras palabras, es necesario aplicar tanto procesos cognitivos relacionados con la creación de significado (e.g., parafrasear o hacer inferencias) como procesos metacognitivos que implican monitorización (e.g., comparar las respuestas propias con el estándar proporcionado) y autorregulación (e.g., cómo resolver problemas específicos). Todo ello permitirá al estudiante confirmar su modelo de respuesta o producir cambios en la representación mental construida sobre la información evaluada en la tarea. Ahora bien, su efectividad puede verse obstaculizada si el compromiso de los participantes para procesar la retroalimentación es bajo (Golke et al., 2015) o los mensajes son difíciles de comprender (Carless, 2006).

Investigaciones previas han encontrado que los estudiantes no sólo desean recibir EF tras resolver tareas (especialmente cuando no reciben retroalimentación correctiva) (Pridemore y Klein, 1991, 1995; Waddick, 1994), sino que además buscan retroalimentación activamente cuando completan evaluaciones en las que se les permite autorregular su propio proceso de aprendizaje (Dunlosky y Rawson, 2015). Como señala Waddick (1994), es posible proporcionar retroalimentación KR automáticamente y ofrecer al estudiante la posibilidad de acceder a mensajes EF cuando lo considere oportuno. Hasta la fecha, existe poca investigación que haya examinado el papel que

juega la toma de decisiones en la búsqueda de retroalimentación formativa en entornos digitales de aprendizaje. Por ejemplo, Pridemore y Klein (1991) encontraron que los estudiantes buscan retroalimentación KR y EF cuando tienen la opción de acceder voluntariamente a esa información. A pesar de los beneficios encontrados para la EF, Pridemore y Klein (1991) no observaron diferencias entre aquellos que tenían la opción de acceder a dicha retroalimentación y aquellos que recibían la retroalimentación de forma automática. Estudios más recientes muestran que los estudiantes deciden buscar voluntariamente EF tras recibir retroalimentación KR, aunque los resultados muestran una alta variabilidad entre los participantes (Fox et al., 2014; Schutz, 1993; Timmers y Veldkamp, 2011).

Dunlosky y Rawson (2015) evaluaron cómo autorregulan el proceso de aprendizaje un conjunto de estudiantes universitarios cuando se les permite estudiar los materiales, evaluarse a sí mismos, o evaluar sus respuestas (i.e., comparar su respuesta con la definición correcta a modo de retroalimentación). Además, los autores examinaron la efectividad de este último procedimiento de aprendizaje autorregulado usando evaluaciones con retroalimentación (i.e., *testing with feedback*) en comparación con un procedimiento controlado (i.e., *test-judge-study*) que consistía en una condición criterio en la que los participantes debían proporcionar la definición correcta un número determinado de veces. Los participantes tuvieron que aprender definiciones sobre el desarrollo cognitivo (e.g., "*¿Qué es la etapa sensoriomotora?*") y completar una prueba final una semana después. Los resultados mostraron que los estudiantes que podían autorregular su aprendizaje utilizaban las evaluaciones y la retroalimentación, lo que significa que monitorizaban su progreso de aprendizaje al evaluarse a sí mismos y al comparar su ejecución con la definición correcta del concepto evaluado. En esas situaciones, los participantes buscaban retroalimentación principalmente para evaluar la precisión de sus respuestas en las preguntas respondidas parcial o totalmente correctas, mientras que decidían reestudiar las definiciones cuando no pudieron recordar la definición. Es decir, los estudiantes empleaban la retroalimentación para verificar la precisión de sus respuestas cuando estaban casi seguros de que eran correctas. Aunque los resultados indicaron que los estudiantes hacen uso de la retroalimentación, Dunlosky y Rawson (2015) sugieren que es necesario realizar nuevas investigaciones para establecer el uso de la retroalimentación en otros escenarios de aprendizaje en los que los estudiantes han de aprender conocimientos conceptuales más complejos.

Incluso cuando los estudiantes deciden procesar la retroalimentación, puede que experimenten dificultades para comprender su contenido y no sean capaces de utilizar esa información para mejorar su aprendizaje (Nuthall, 2007). Los trabajos presentados en esta tesis prestan especial atención al uso y procesamiento de la EF en un entorno digital, ya que este tipo de retroalimentación es el más efectivo para promover el aprendizaje (Hattie y Timperley, 2007; Shute, 2008; Van der Kleij et al., 2015). Además, la EF suele incluir retroalimentación correctiva KR o KCR, la cual puede determinar cómo emplean la EF los estudiantes. En entornos académicos, responder preguntas de comprensión a partir de textos de lengua o ciencias requiere generalmente un procesamiento de orden superior, ya que el estudiante ha de hacer inferencias entre conceptos e incluso utilizar la información aprendida para resolver tareas distintas. Estas situaciones requieren que los profesores o los programas informáticos proporcionen retroalimentación con explicaciones o pautas metacognitivas para mejorar el aprendizaje de los estudiantes. Procesar esta retroalimentación en el contexto de la tarea puede considerarse una actividad compleja en sí misma. Por ejemplo, recibir información sobre por qué la temperatura y la altitud afectan a la presión atmosférica requiere de la aplicación de procesos cognitivos y metacognitivos que permitan al estudiante comprender los conceptos y relaciones entre conceptos necesarios para corregir o añadir a su representación mental el efecto de ambos factores sobre la presión atmosférica.

A pesar de la vasta literatura sobre la efectividad de la retroalimentación, aún se desconocen los mecanismos específicos que relacionan el procesamiento de la retroalimentación con el aprendizaje. A lo largo de las últimas décadas se han propuesto varios modelos sobre retroalimentación, entre los que destacan el modelo de retroalimentación de tres ciclos (Kulhavy y Stock, 1989), el modelo de cinco etapas para recibir retroalimentación (Bangert-Drowns et al., 1991), el modelo de aprendizaje autorregulado (Butler y Winne, 1995), la teoría de la intervención de la retroalimentación (Kluger y DeNisi, 1996), un modelo de aprendizaje autorregulado con retroalimentación en evaluaciones formativas (Nicol y Macfarlane-Dick, 2006), un modelo de retroalimentación para facilitar el aprendizaje (Hattie y Timperley, 2007), el modelo de retroalimentación tutorizada interactiva (modelo ITF; Narciss, 2013), o el modelo de procesamiento de la retroalimentación en entornos inteligentes de aprendizaje (modelo LP-FILE; Timms, DeVelle y Lay, 2016). Este conjunto de teorías cuenta con la retroalimentación como un componente central en el proceso de

aprendizaje. A continuación explicamos brevemente las principales aportaciones de cada una de ellas al objeto de clarificar nuestro marco de trabajo, desde el que intentamos comprender los procesos (meta)cognitivos y el proceso de toma de decisiones al recibir retroalimentación formativa en un entorno digital.

Kulhavy y Stock (1989). Estos autores propusieron un modelo que tenía en cuenta el papel de la retroalimentación en la corrección de errores. Los autores sugirieron que las expectativas que generan los estudiantes acerca de sus respuestas influyen en la forma en que hacen uso de la retroalimentación, por lo que el grado de atención dedicado a la retroalimentación varía en función del nivel de confianza en sus respuestas. Cuando se proporciona retroalimentación sobre la precisión de las respuestas, el nivel de confianza en la respuesta interactúa con la retroalimentación, produciendo o no una discrepancia entre el rendimiento actual y el deseado. Kulhavy y Stock (1989) encontraron que la retroalimentación juega un papel importante cuando los estudiantes responden incorrectamente pero tienen un nivel alto de confianza en sus respuestas. En esta situación, los estudiantes tienden a dedicar más tiempo a procesar la retroalimentación en comparación con las situaciones en las que tienen un nivel bajo de confianza. Si el nivel de confianza es bajo, los estudiantes dedican menos recursos a procesar la retroalimentación, ya que esta situación no estimula la reflexión y el consiguiente aprendizaje. Asimismo, los estudiantes dedican poco tiempo a procesar la retroalimentación cuando responden correctamente y están seguros de su respuesta. Estos resultados sugieren que los estudiantes generan objetivos y monitorizan el uso de la retroalimentación en función de esos objetivos y la discrepancia entre la seguridad en sus respuestas y la exactitud de las mismas. Sin embargo, el modelo propuesto por Kulhavy y Stock no contempla al aprendiz como un agente activo.

Bangert-Drowns et al. (1991). Estos autores formularon un modelo que describe los efectos de la retroalimentación en los estados cognitivos del aprendiz en cinco fases: estado inicial, búsqueda/recuperación de información, respuesta, evaluación y ajuste. Al comienzo, el estudiante cuenta con un estado inicial formado por una serie de características (Fase 1) como, por ejemplo, su interés, sus metas, su percepción de autoeficacia, y su nivel de conocimiento previo. Cuando se presenta una tarea, el estudiante debe activar estrategias de búsqueda o recuperación de información para resolverla (Fase 2). Una vez que ha localizado la información necesaria, el estudiante proporciona una respuesta con cierto grado de confianza, generando así expectativas sobre la retroalimentación que recibirá (Fase 3). Al recibir la retroalimentación, el

estudiante debe evaluar su rendimiento en función de la información proporcionada (Fase 4), por lo que esta evaluación puede variar según su nivel de confianza en la respuesta y el éxito al responder (Kulhavy, 1977; Kulhavy y Stock, 1989). Los estudiantes utilizan la retroalimentación para verificar la precisión de sus respuestas y corregir sus errores. Dicha evaluación realizada por el estudiante al recibir retroalimentación da lugar a una serie de ajustes o modificaciones en sus conocimientos, intereses, objetivos y nivel de autoeficacia (Fase 5), los cuales pasarán a formar parte del nuevo estado inicial cuando complete nuevas tareas. Aunque los autores señalan que desconocen los procesos cognitivos involucrados en la evaluación de las respuestas, afirman que la retroalimentación debe ser recibida con atención para producir efectos en el aprendizaje: “*feedback can promote learning if it is received mindfully*” (Bangert-Drowns et al., 1991).

Butler y Winne (1995). Estos autores propusieron un modelo en el que la retroalimentación es un componente central de cualquier escenario de aprendizaje autorregulado. En general, el aprendizaje autorregulado se rige por un proceso recursivo, en el cual los aprendices monitorizan internamente su estado al completar una tarea, autorregulan su ejecución, y reciben retroalimentación que sirve para mejorar la regulación en futuras tareas. Estos autores argumentaron acerca de la relación entre la retroalimentación y los procesos cognitivos involucrados en tareas que requieren autorregulación. Cuando un estudiante realiza una tarea de aprendizaje, construye una representación de las características y requisitos de la misma basándose en sus conocimientos y creencias. A partir de esa representación, los aprendices establecen objetivos que intentan alcanzar mediante la aplicación de estrategias que dan lugar a productos de tipo mental y conductual. Este conjunto de procesos activados y productos generados durante la resolución de la tarea promueve la formación de retroalimentación interna (i.e., el aprendiz genera metas, aplica estrategias, y monitoriza su ejecución). Esta información puede influir en el aprendiz, quien tal vez modifique sus objetivos, adapte sus estrategias y habilidades, y genere nuevos procedimientos para resolver la tarea. Posteriormente, si el estudiante recibe retroalimentación externa, esta información le permite establecer comparaciones entre su retroalimentación interna y el estándar proporcionado. Así, pueden existir coincidencias o discrepancias entre el rendimiento alcanzado y el deseado. En caso de existir discrepancias, el estudiante debe iniciar un proceso de revisión mediante el que modificar sus conocimientos y creencias, mejorando su autorregulación. Siguiendo el trabajo de Chinn y Brewer (1993) sobre

cómo los conceptos erróneos pueden obstaculizar el cambio conceptual, Butler y Winne (1995) expusieron seis formas desadaptadas de búsqueda y procesamiento de la retroalimentación externa. Así, los estudiantes pueden a) ignorar la retroalimentación, b) rechazarla, c) considerarla irrelevante, d) evaluarla como no relacionada, e) reinterpretarla para que se ajuste a la retroalimentación interna, o f) realizar cambios superficiales en sus conocimientos y/o creencias a partir de esa información. A pesar de la relación establecida entre la retroalimentación y los procesos involucrados en el aprendizaje autorregulado, Butler y Winne (1995) no explican qué operaciones mentales se producen cuando el estudiante recibe retroalimentación externa, aspecto esencial para comprender sus efectos en el proceso de monitorización durante la realización de tareas que requieren alto nivel de autorregulación.

Kluger y DeNisi (1996). Estos autores propusieron una teoría sobre el papel de la retroalimentación como intervención, prestando atención a la influencia que tienen otras variables moderadoras en su efectividad como, por ejemplo, el formato de presentación de la retroalimentación o las características específicas de la tarea. El supuesto principal sobre el que se construye esta teoría es que las intervenciones con retroalimentación dirigen el locus de atención del aprendiz a alguno de estos tres niveles: (a) procesos de aprendizaje de la tarea, (b) procesos de motivación por la tarea, y (c) procesos de meta-tarea que incluyen información del *self*. La retroalimentación que dirige el locus de atención hacia aspectos de la tarea mejora el aprendizaje en comparación con la retroalimentación que se centra en procesos de meta-tarea relacionados con el *self*. No obstante, los resultados encontrados en su meta-análisis muestran una alta variabilidad en cuanto a la efectividad de la retroalimentación, por lo que los autores sugieren que debe explorarse en detalle qué retroalimentación promueve significativamente el aprendizaje y bajo qué circunstancias. Kluger y DeNisi (1996) propusieron cinco argumentos interrelacionados sobre los que se fundamenta su teoría: (a) el comportamiento está regulado por comparaciones entre la retroalimentación con los objetivos o estándares, (b) los objetivos o estándares están organizados jerárquicamente, (c) la atención es limitada, por lo que sólo las discrepancias entre el desempeño real y el deseado que reciben atención modifican la regulación del comportamiento, (d) la atención se dirige normalmente a un nivel moderado de la jerarquía, y (e) la retroalimentación cambia el locus de atención del aprendiz y, por tanto, influye en su comportamiento. Como señalan los autores, este último aspecto es esencial para comprender la relación retroalimentación-aprendizaje. Así, Kluger y DeNisi consideran

que el rendimiento y la motivación son efectos de la retroalimentación, y señalan que lo verdaderamente importante es conocer la reacción a esa información, es decir, los procesos inducidos por la presencia de la retroalimentación. A pesar de su reconocida importancia, los autores no indagaron en las operaciones mentales y la toma de decisiones que realizan los estudiantes cuando reciben retroalimentación formativa.

Nicol y Macfarlane-Dick (2006). Basándose en el modelo inicialmente propuesto por Butler y Winne (1995), estos autores desarrollaron un modelo de aprendizaje autorregulado que incluía retroalimentación en evaluaciones formativas. Aunque estos autores desarrollaron su teoría centrándose en estudiantes de educación superior, la propuesta de Nicol y Macfarlane-Dick contiene ideas valiosas para el aprendizaje autorregulado en cualquier etapa educativa. Este modelo asume que los estudiantes han de participar activamente en el procesamiento y la comprensión de la retroalimentación externa (Black y Wiliam, 1998), así como en la monitorización y autorregulación de su propio aprendizaje (Butler y Winne, 1995). Según el modelo, el profesor asigna una tarea (A) que desencadena una serie de procesos de autorregulación en el estudiante. Así, la tarea requiere que el estudiante recurra a sus conocimientos previos y motivaciones (B), y construya una interpretación personal de la tarea y sus requisitos. En base a esta interpretación, el estudiante formula sus propios objetivos de la tarea (C). Aunque generalmente se espera una superposición entre los objetivos del alumno y los del profesor, el grado de superposición puede variar y no siempre ser compartido. Esos objetivos determinan las estrategias y tácticas que emplean los estudiantes (D) para generar resultados, tanto internos (i.e., cambios en los estados cognitivos o afectivo-motivacionales que ocurren durante la tarea) (E) como externos (i.e., productos tangibles generados o comportamientos) (F). El proceso de monitorización durante la tarea (i.e., comparaciones del progreso con los objetivos deseados) genera retroalimentación interna a nivel cognitivo, motivacional y conductual. Esas comparaciones ayudan al estudiante a determinar si la forma en que se está realizando la tarea debe continuar igual o es necesario hacer algún ajuste. Cuando el estudiante recibe retroalimentación externa (G) sobre sus resultados externos, la retroalimentación puede aumentar, coincidir o entrar en conflicto con la interpretación de la tarea y el proceso de aprendizaje del estudiante. Para que esta retroalimentación produzca un efecto en los procesos internos o en los resultados externos, el estudiante debe interpretar esa información activamente. Por tanto, el modelo asume que la retroalimentación, bien sea interna o externa, permite al estudiante conocer la relación

entre el nivel de aprendizaje alcanzado y los estándares externos. A pesar de que Nicol y Macfarlane-Dick (2006) expusieron una serie de principios para el desarrollo de la retroalimentación que facilitan el aprendizaje autorregulado, los autores no explicaron qué significa interpretar la retroalimentación externa en el contexto de la tarea de aprendizaje, un elemento esencial para comprender cómo y porqué la retroalimentación puede influir en las representaciones mentales de los estudiantes.

Hattie y Timperley (2007). Estos autores propusieron un modelo bajo la premisa de que el principal objetivo de la retroalimentación es reducir la discrepancia entre la comprensión actual y la comprensión deseada, para lo cual los estudiantes aplican estrategias que influyen en su aprendizaje. Para reducir esta discrepancia, los estudiantes pueden dedicar más esfuerzos o emplear estrategias más efectivas (e.g., desarrollar habilidades para detectar errores) hasta reducir o incluso abandonar sus objetivos. Igualmente, los profesores pueden reducir esta discrepancia mediante el uso de objetivos específicos o proporcionando retroalimentación para desarrollar estrategias. Según propusieron Hattie y Timperley (2007), para que la retroalimentación sea efectiva ha de responder a tres cuestiones: a) ¿Cuáles son los objetivos?, b) ¿Qué progreso estoy haciendo para alcanzarlos?, y c) ¿Qué necesito hacer a continuación para mejorar? La efectividad a la hora de responder estas preguntas depende del nivel al que se dirige la retroalimentación. Así, la retroalimentación puede situarse en el nivel de tarea, relacionado con el rendimiento obtenido; en el nivel de proceso, relacionado con el proceso de comprensión para completar la tarea; en el nivel de autorregulación, relacionado con el proceso de monitorización y control regulatorio de las acciones; o en el nivel personal/self, relacionado con evaluaciones del propio aprendiz. Mientras que este último nivel es el menos efectivo, los niveles de proceso y autorregulación son los que guardan una relación más directa con la efectividad de la retroalimentación, ya que ofrecen la posibilidad de mejorar la comprensión y el proceso de control durante la realización de una tarea. Aunque la perspectiva de estos autores abre nuevos caminos para comprender algunos componentes de la retroalimentación en un entorno educativo tradicional, no proporciona información sobre cómo es recibida la retroalimentación por el estudiante, quien debe intentar reducir la discrepancia entre su comprensión actual y la deseada.

Narciss (2013). Esta autora desarrolló el modelo de retroalimentación tutorizada interactiva (*Interactive Tutoring Feedback o ITF*), un modelo que conceptualiza la retroalimentación como una actividad instruccional multidimensional, cuyo fin último

es contribuir a la regulación del proceso de aprendizaje de forma que los estudiantes adquieran el conocimiento y las competencias necesarias para realizar las tareas de aprendizaje. Narciss considera que, en un contexto instructivo en el que se proporciona retroalimentación externa, se deben tener en cuenta dos ciclos de retroalimentación interactivos: el ciclo de retroalimentación interno del estudiante y el ciclo de retroalimentación externo proporcionado por un recurso externo. La interacción entre ambos ciclos de retroalimentación se produce gracias al proceso controlado (*controlled process*), encargado de la adquisición de competencias (i.e., conocimientos y estrategias) relacionadas con las tareas de aprendizaje. Según Narciss, las competencias han de ser controladas a nivel cognitivo, metacognitivo y motivacional, por lo que es esencial identificar qué variables son relevantes para la adquisición de cada competencia. Para cada variable, el estudiante genera un estándar, es decir, una representación subjetiva de la competencia a desarrollar que puede estar influida por características individuales, como los objetivos establecidos o el nivel de conocimiento previo. Al mismo tiempo, existe un estándar externo definido por una representación externa de la competencia, que puede estar influida por factores instruccionales, como los objetivos de la tarea o el contenido de la misma. Los ciclos de retroalimentación requieren que el aprendiz y la fuente de retroalimentación externa realicen una evaluación del estado de las competencias, dando lugar a una retroalimentación interna y externa de las mismas. Tanto la retroalimentación interna como la externa necesitan de un controlador interno y externo, es decir, una unidad de procesamiento de información que regule el proceso de adquisición de competencias. Estos controladores permiten establecer comparaciones entre la ejecución del estudiante y el estándar establecido. Si existen discrepancias entre el nivel actual y el nivel estándar de competencia, el sistema proporcionará EF para reducir la distancia entre los niveles. Esta retroalimentación externa ha de ser procesada por el controlador interno del estudiante junto con la retroalimentación interna, estableciendo así comparaciones entre los estándares y su ejecución. Estas comparaciones activan un proceso de control interno (*internal control action*) basado en las discrepancias entre la retroalimentación interna y externa, a partir de las cuales los estudiantes generan ideas sobre cómo mejorar la competencia aprendida. Por último, el proceso de control interno envía información al proceso controlado, donde se implementan acciones para reducir las discrepancias entre los niveles actual y deseado.

Timms et al. (2016). Estos autores han propuesto el modelo de procesamiento de la retroalimentación en entornos inteligentes de aprendizaje (LP-FILE model), un modelo general de toma de decisiones que captura los pasos en los que los alumnos perciben, decodifican, reciben y dan significado a la información proporcionada. De acuerdo con el modelo LP-FILE, los estudiantes deben cometer un error para iniciar conductas de revisión. En ese momento, el estudiante puede detectar el error y corregirlo gracias a su sistema de monitorización interno. En caso de que el estudiante no detecte el error, el sistema informático rastrea el error y envía retroalimentación externa al estudiante, quien debe percibirlo y decodificar su contenido. Una vez decodificada la retroalimentación, el estudiante debe comprender el significado del mensaje de retroalimentación en el contexto de la tarea realizada. Si este proceso de retroalimentación funciona adecuadamente, el estudiante podría corregir su error inicial. Aunque *Timms et al. (2016)* describieron los pasos secuenciales que realizan los estudiantes cuando reciben retroalimentación en un entorno digital de aprendizaje, el modelo no contempla las operaciones cognitivas y metacognitivas que los estudiantes implementan durante su procesamiento. De hecho, como sugieren los autores, recoger información sobre cómo decodifican y comprenden los mensajes de retroalimentación los estudiantes es una tarea compleja porque requiere de procesos mentales internos. Para ello, el uso de protocolos de pensamiento en voz alta puede ser una herramienta útil para capturar los procesos mentales que aparecen durante el procesamiento de la retroalimentación (*Krause, Stark y Mandl, 2009*). Además, es importante señalar que los dos últimos modelos aquí presentados centran su atención en el procesamiento de la retroalimentación en un entorno digital.

Este corpus de trabajos empíricos y teóricos sugiere que la retroalimentación puede considerarse como un evento único de aprendizaje que es necesario comprender en detalle. La mayoría de los modelos presentados anteriormente consideran que el estudiante ha de tomar el control de la retroalimentación y participar activamente en su procesamiento. Además, investigaciones previas demuestran que los aprendices buscan retroalimentación acerca de su rendimiento a menudo, incluso cuando esta información no afecta a su aprendizaje (*Kluger y Adler, 1993; Kluger, Adler y Fay, 1992*). Asimismo, parece que la efectividad de las tareas asignadas depende de si los estudiantes están dispuestos a buscar activamente retroalimentación y construir significado a partir de su contenido (e.g., *Bangert-Drowns et al., 1991; Nicol y Macfarlane-Dick, 2006*). Ahora bien, ¿qué significa en términos cognitivos y

metacognitivos participar activamente en el procesamiento de la retroalimentación? ¿En qué medida están dispuestos los estudiantes a buscar EF en un entorno digital de aprendizaje? Basándonos en investigaciones teóricas y empíricas previas, hemos desarrollado un marco de trabajo sobre el procesamiento de la retroalimentación que tiene en cuenta los cambios en el estado del modelo de respuesta que generan los estudiantes cuando responden preguntas de un texto.

El modelo presentado en la Figura 2 parte de la siguiente situación: cuando se responde una pregunta y se recibe EF, ofrecida generalmente junto con retroalimentación correctiva, el estudiante debe decidir si es necesario revisar su modelo de respuesta a la luz de la retroalimentación proporcionada. Al responder una pregunta, puede que los estudiantes no hayan integrado el nuevo conocimiento con sus conocimientos previos debido a varias causas como, por ejemplo, la activación de conocimientos previos erróneos, dificultades para localizar información relevante, o limitaciones para comprender el contenido del texto o de la pregunta. Esta situación requiere que el estudiante revise su conocimiento a partir de la información proporcionada en la retroalimentación.

El modelo tiene en cuenta dos fases al recibir retroalimentación: Fase 1, Verificación de respuestas, y Fase 2, Revisión de conocimiento. Al responder cada pregunta, el estudiante formula, con diferente nivel de confianza, un modelo de respuesta inicial más o menos preciso (Estado 1). Una vez que el estudiante valida su respuesta, se inicia la primera fase (i.e., Verificación de respuestas), en la que el estudiante quiere conocer si su respuesta es correcta o no. Si la respuesta es correcta, la retroalimentación, especialmente la de tipo correctivo, permite al estudiante verificar su modelo de respuesta (Estado 2a) y aumentar así su nivel de seguridad en sus conocimientos. No obstante, si la respuesta es errónea, la retroalimentación de tipo correctivo refuta el modelo de respuesta del estudiante (Estado 2b). Cuando se proporciona EF, el estudiante ha de decidir si necesita o quiere procesar retroalimentación adicional después de haber verificado o refutado su modelo de respuesta inicial. Esta segunda fase (i.e., Revisión de conocimiento) se iniciará principalmente cuando el estudiante proporcione respuestas incorrectas (i.e., el modelo de respuesta inicialmente construido es refutado) o respuestas correctas con inseguridad (i.e., el estudiante no acaba de comprender el conocimiento evaluado o simplemente ha proporcionado la respuesta correcta por mera probabilidad). Así, después de verificar o refutar su modelo de respuesta, el estudiante puede omitir procesar la retroalimentación

adicional por diversas razones (e.g., ha proporcionado la respuesta correcta y considera innecesaria la información adicional, está interesado únicamente en conocer la precisión de sus respuestas, o no dispone de tiempo). En este caso, asumimos que no se producirán modificaciones relevantes en su representación mental de los conocimientos evaluados. Sin embargo, si el estudiante decide procesar la retroalimentación, entonces inicia un proceso de revisión de sus conocimientos que implica operaciones cognitivas y metacognitivas que favorezcan la reestructuración de conocimientos imprecisos o erróneos, dando lugar a actualizaciones en el modelo de respuesta (Estado 3). En ese momento, el estudiante ha de establecer comparaciones entre el modelo de respuesta activo en su memoria de trabajo y la retroalimentación, información que compite con el conocimiento incompleto o incorrecto previamente adquirido. Cuando intenta comprender la retroalimentación, el estudiante realiza esfuerzos por añadir información o corregir su representación mental evaluada en la tarea. Así, el procesamiento de la retroalimentación conlleva la aplicación de procesos cognitivos (i.e., construcción de significado) y metacognitivos (i.e., monitorización y autorregulación). Asumimos que las actualizaciones del modelo de respuesta pueden variar ampliamente entre estudiantes e incluso entre preguntas, ya que esas modificaciones dependen de la disposición y habilidad del estudiante para comprender la retroalimentación en el contexto de tarea. Este procesamiento puede abarcar desde simples modificaciones en el modelo de respuesta (e.g., “*X es la respuesta correcta*”, “*Los factores externos afectan a la presión atmosférica*”) a modificaciones más profundas que requieran de procesos elaborativos más complejos (e.g., “*La presión atmosférica puede variar en función de factores externos como la temperatura o la altitud*”). A pesar de que se produzcan modificaciones en el conocimiento previamente adquirido, no podemos asumir que los estudiantes lograrán actualizar sus modelos de respuesta adecuadamente, incluso cuando la EF sea procesada activamente. Finalmente, el estudiante deberá decidir cuándo detener el procesamiento y continuar con la siguiente tarea. Así, mientras el estudiante no tome la decisión de salir de la retroalimentación, podemos asumir que seguirá involucrado en la revisión de sus conocimientos. Cuando finalice su procesamiento, entonces continuará con la tarea y pasará a responder una nueva pregunta.

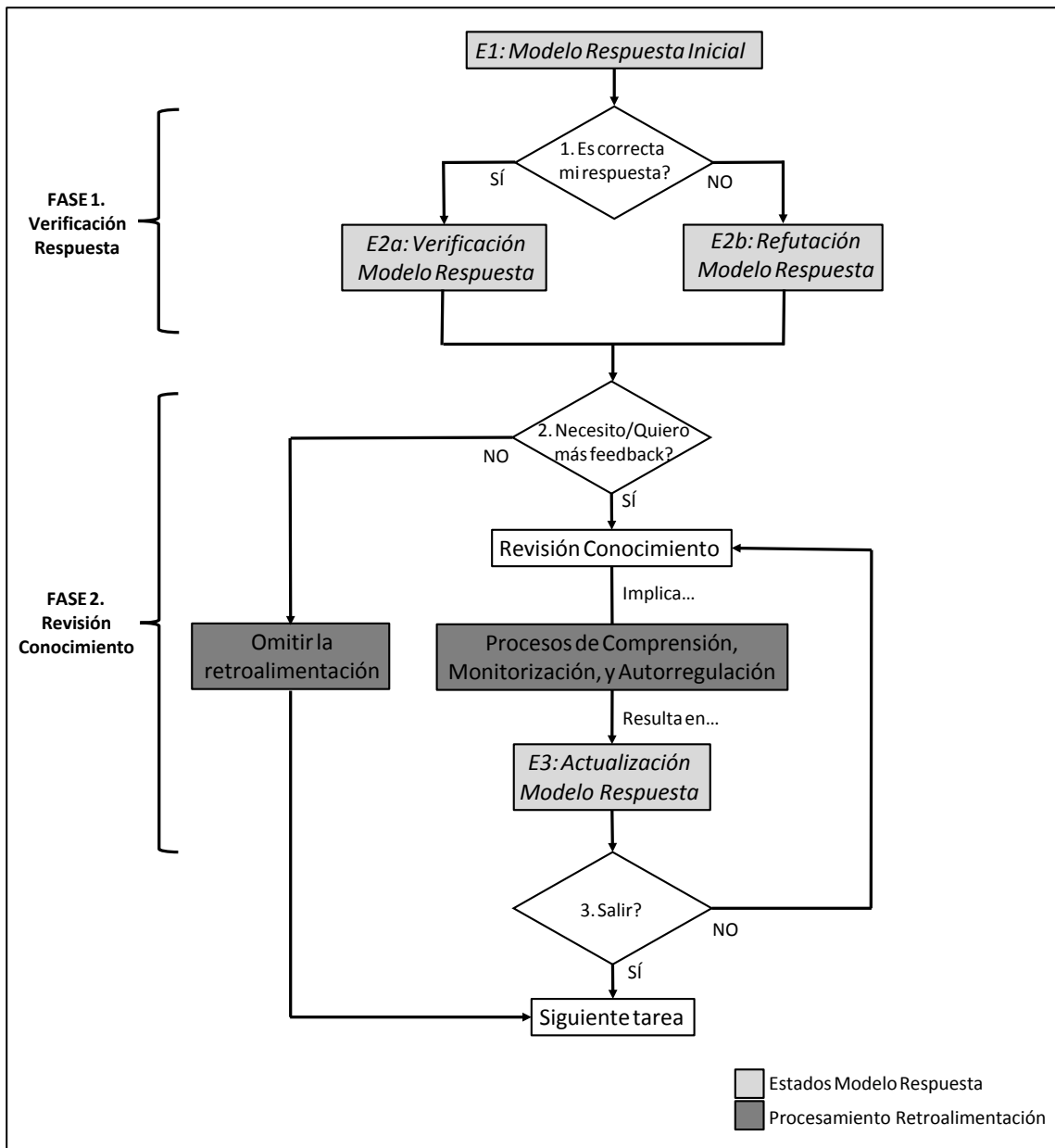


Figura 2. Esquema de procesamiento de la EF al responder preguntas de un texto.

Además, asumimos que factores relacionados con la propia retroalimentación (e.g., la presencia de retroalimentación correctiva), las características individuales de los estudiantes (e.g., competencia lectora o conocimiento previo), las tareas (e.g., dificultad de las preguntas) o características contextuales (e.g., disponibilidad de los materiales) pueden influir en el procesamiento de la retroalimentación formativa en entornos digitales de aprendizaje. Una de las principales variables que puede influir en el procesamiento de la EF es la presencia de retroalimentación correctiva KR o KCR. En caso de no recibir retroalimentación correctiva sobre su ejecución, el alumno deberá tomar la decisión de procesar la EF en función de su evaluación interna acerca de la

precisión de su respuesta. Desafortunadamente, estas evaluaciones suelen ser imprecisas y puede que afecten al uso de la EF en función de su rendimiento real. En caso de recibir retroalimentación correctiva, la decisión de utilizar la EF puede verse modificada por una evaluación externa sobre la precisión de su respuesta. Al conocer que una respuesta es incorrecta, la retroalimentación desencadena una discrepancia cognitiva que debe corregirse. En ese caso, los estudiantes pueden usar la EF para mejorar su nivel de comprensión y, por tanto, su aprendizaje de los conocimientos evaluados en la tarea. Aunque la retroalimentación correctiva KR y KCR mejore el proceso de monitorización de los estudiantes, aumentando el uso de la EF tras responder incorrectamente, puede que esta información correctiva, especialmente la KCR, inhiba el procesamiento de la EF, ya que cualquier información adicional puede considerarse innecesaria.

Es importante señalar que la fase de Verificación consume pocos recursos cognitivos, ya que el objetivo principal es validar o refutar el modelo de respuesta inicialmente construido. Sin embargo, la fase de Revisión de conocimiento requiere generalmente altos niveles de procesamiento, ya que el estudiante necesita aplicar operaciones cognitivas y metacognitivas que le permitan generar cambios en su representación mental de los conocimientos evaluados. A pesar de la importancia dada al procesamiento activo de la retroalimentación para que se produzcan modificaciones cognitivas en la representación mental de los estudiantes, existe escasa investigación sobre el uso y procesamiento de la EF en entornos digitales. Mientras que los procesos cognitivos se refieren a los procesos de comprensión relacionados con operaciones como leer, parafrasear o elaborar la información proporcionada en la retroalimentación, los procesos metacognitivos se refieren a la monitorización y autorregulación que hacen los estudiantes de esa información. Aunque estos procesos pueden variar según el tipo de EF recibida, nuestro objetivo es proporcionar una perspectiva general sobre los principales procesos (meta)cognitivos que los estudiantes pueden implementar al procesar la retroalimentación formativa en un entorno digital.

Procesos de Comprensión

La retroalimentación suele proporcionarse como un mensaje escrito que el estudiante ha de leer, por lo que necesita activar y organizar información para comprender el contenido del texto (McNamara y Magliano, 2009). Aunque los modelos clásicos de comprensión conceptualizan esta capacidad como la construcción de una

representación mental coherente a partir de la información del texto (Kintsch, 1998), comprender la información proporcionada en la retroalimentación formativa va más allá de esta idea. Según Butler y Winne (1995), la retroalimentación que se proporciona al alumno sirve para "*confirmar, sobrescribir, agregar, ajustar o reestructurar los conocimientos y creencias existentes*". Estos propósitos reconocen el hecho de que los estudiantes deben realizar procesos de orden superior, ya que la retroalimentación generalmente se proporciona para corregir ideas incorrectas. Para que la retroalimentación sea efectiva, los estudiantes deben participar activamente en su procesamiento y ser capaces de comprender su contenido. Por tanto, procesar la retroalimentación es una cuestión de comprender la información y revisar los conocimientos adquiridos. Así, el estudiante tiene que integrar la información de la retroalimentación en su representación mental, a menudo incompleta o incorrecta. En este proceso, el alumno tiene que activar ideas de su conocimiento previo para hacer inferencias, algunas de las cuales pueden ser inexactas o incorrectas, por lo que pueden impedir que los estudiantes actualicen su modelo de respuesta y, en consecuencia, construyan una representación mental coherente e integrada de los conocimientos que han de aprender. La retroalimentación requiere no sólo activar conocimiento previo, sino también desvincularse de él y revisarlo para alcanzar una comprensión adecuada.

Se asume que los estudiantes implementan estos procesos cognitivos durante el procesamiento de la retroalimentación, aunque se desconoce hasta qué punto esto es así. Estos procesos cognitivos difieren según su complejidad. En algunos casos, los estudiantes simplemente leen la retroalimentación y crean un significado literal de su contenido. En otros casos, los estudiantes participan de una manera más activa y parafrasean o elaboran a partir de la información. La retroalimentación permite a los estudiantes contrastar su rendimiento actual con el nivel deseado, lo que implica conectar el contenido de la retroalimentación con su conocimiento adquirido por medio de inferencias para comprender la respuesta correcta o la razón por la cual su respuesta es incorrecta. Estos procesos inferenciales implican un procesamiento de orden superior que pueda mejorar el aprendizaje profundo. Cuando el estudiante decide participar activamente en el procesamiento de la retroalimentación, se espera que revise su conocimiento actual (i.e., el modelo de respuesta incompleto o incorrecto) para integrar nueva información en su representación mental. Por tanto, los procesos de elaboración en la memoria de trabajo del estudiante son necesarios para alcanzar cierto nivel de comprensión que establezca el conocimiento codificado en su representación mental.

En resumen, los estudiantes deben procesar la retroalimentación formativa y construir significado al comparar su modelo de respuesta inicial con el estándar habitualmente proporcionado. Si el alumno activa los recursos cognitivos necesarios para procesar el contenido de la retroalimentación, es posible que integre esta información con su conocimiento existente, estableciendo nuevas relaciones entre la información de la retroalimentación recientemente codificada y el conocimiento previamente adquirido. El resultado de este procesamiento debería conducir a mejores representaciones mentales. En definitiva, la EF debe interpretarse a la luz de la tarea realizada para lograr ganancias en el aprendizaje (Timms et al., 2016).

Procesos de Monitorización

La naturaleza de la retroalimentación es desafiar la cognición del alumno cuando está equivocado y estabilizar su conocimiento cuando está en lo cierto. Como sugieren Butler y Winne (1995), la retroalimentación puede iniciar procesos de monitorización que evalúan a nivel superficial la información recién aprendida, especialmente cuando la retroalimentación correctiva KR y KCR informa de respuestas correctas. Sin embargo, la EF puede iniciar procesos de monitorización que mejoren tanto los procesos cognitivos implicados en la tarea como los resultados obtenidos. Por tanto, los estudiantes deben monitorizar el procesamiento de la retroalimentación, iniciando una serie de evaluaciones sobre su rendimiento y/o sobre ellos mismos.

En primer lugar, los estudiantes deben controlar el uso de la retroalimentación al evaluar la precisión de sus respuestas con respecto al nivel deseado de rendimiento. Ese proceso implica hacer comparaciones entre su modelo de respuesta y el estándar proporcionado en la retroalimentación. Esto permite a los estudiantes verificar su modelo de respuesta y evaluar la necesidad de actualizarlo de acuerdo con la información proporcionada en la retroalimentación. Si eso sucede, es posible que se den cambios cognitivos en las representaciones mentales de los estudiantes. En segundo lugar, los estudiantes también pueden usar los mensajes para evaluar su grado de comprensión o aprendizaje; esto es, el proceso de supervisión y validación consciente de la propia comprensión. Además de estos procesos, investigaciones recientes sugieren que los estudiantes que reciben retroalimentación realizan dos juicios motivacionales, uno atribucional al recibir la retroalimentación y uno de valoración de cara a la siguiente tarea (e.g., Crombach, Boekaerts y Voeten, 2003; Timmers, Braber-Van Den Broek y Van Den Berg, 2013), por lo que la retroalimentación puede activar procesos de tipo

afectivo-motivacional. Así, los estudiantes pueden reaccionar afectivamente ante la presencia de retroalimentación (Kluger y DeNisi, 1996), especialmente la de carácter correctivo. La retroalimentación externa puede influir en cómo se sienten los alumnos, tanto positiva como negativamente (Dweck, 1999). De hecho, es posible que el efecto más inmediato de la retroalimentación relacionado con el rendimiento sea una respuesta afectiva (Pekrun, 2006). Los estudiantes pueden experimentar alegría, alivio, tristeza, enojo o incluso desesperación al recibir los mensajes de retroalimentación. Por último, la retroalimentación puede activar procesos atribucionales cuando los estudiantes evalúan la causa por la cual cometieron un error o tuvieron éxito al resolver una tarea. En general, este conjunto de evaluaciones permite a los estudiantes hacer ajustes en su proceso de aprendizaje durante el procesamiento de la retroalimentación.

Procesos de Autorregulación

El uso de entornos digitales de aprendizaje requiere que los estudiantes implementen procesos metacognitivos continuamente, lo que significa que los estudiantes necesitan monitorizar y autorregular el uso de los materiales de aprendizaje. Estos escenarios han revelado que los estudiantes a veces no logran regular exitosamente su proceso de aprendizaje (Lee et al., 2010). Los estudiantes toman decisiones metacognitivas de forma imprecisa habitualmente (e.g., Azevedo y Cromley, 2004), por lo que puede que también autorregulen de forma poco precisa el uso de la EF. Mientras que algunos estudiantes tal vez orienten su atención hacia los componentes correctivos de la retroalimentación KR o KCR, otros puede que busquen EF (e.g., explicaciones, sugerencias o pistas) para comprender mejor por qué una respuesta específica es correcta o incorrecta. En esta línea, los estudiantes han de tomar decisiones acerca de qué información necesitan/quieren procesar. Esta toma de decisiones es el resultado de interacciones complejas entre factores internos y externos. Siguiendo la teoría del aprendizaje autorregulado, los estudiantes utilizan la monitorización del progreso para ajustar sus comportamientos y alcanzar sus objetivos (Winne y Hadwin, 1998). Por lo tanto, podemos decir que los estudiantes tienen que regular el uso de la retroalimentación de forma apropiada, es decir, han de tomar decisiones sobre qué información procesar y durante cuánto tiempo.

2.3.1. Supuestos del procesamiento de la retroalimentación

Tras años de investigación, las perspectivas propuestas por las teorías conductuales que afirman que la retroalimentación funciona como un refuerzo han recibido escaso apoyo, ya que algunos de sus principios no siempre se aplican a las funciones de la retroalimentación, especialmente cuando se refiere al aprendizaje de materiales complejos en entornos educativos (Kulhavy, 1977). Recientemente, diferentes teorías del procesamiento de la retroalimentación han tratado de describir la secuencia de pasos que los estudiantes realizan cuando se proporciona retroalimentación, señalando que los estudiantes tienen que dar significado a los mensajes de retroalimentación en el contexto de la tarea realizada (Timms et al., 2016). Sin embargo, "dar significado" a la retroalimentación, especialmente cuando se refiere a la EF, es una actividad compleja que implica procesos cognitivos y metacognitivos que pueden determinar su efectividad. Para que la retroalimentación formativa produzca efectos en el aprendizaje, asumimos que deben cumplirse una serie de supuestos, entre los que tomamos y adaptamos al campo de la retroalimentación los supuestos 1-5 propuestos en el modelo RESOLV, desarrollado por Rouet et al. (2017), quienes definen la lectura como una actividad dirigida a resolver problemas. Consideramos que estos supuestos representan perfectamente los criterios necesarios para que el procesamiento de la retroalimentación afecte al aprendizaje de los estudiantes.

1. *Procesamiento activo.* Los estudiantes deben estar dispuestos a participar activamente en el procesamiento de la retroalimentación. Aunque algunos investigadores y profesionales parecen asumir que la retroalimentación se procesa automáticamente cuando se proporciona, la realidad es que procesar esos mensajes depende de una decisión voluntaria del alumno, especialmente en entornos digitales en los que la autorregulación del propio aprendizaje cobra especial relevancia. Esto es, los alumnos autorregulan el uso de la retroalimentación y, en consecuencia, deciden en qué medida deben invertir recursos cognitivos para procesar esa información. Una vez que los estudiantes deciden invertir esfuerzos en procesar la retroalimentación, estos han de ser capaces de interpretar su contenido en el contexto de la tarea realizada. Aunque la retroalimentación correctiva no requiere grandes cantidades de atención o habilidades de procesamiento para comprender su información, la EF requiere conocimientos y habilidades para lograr una interpretación precisa que permita al estudiante mejorar su aprendizaje. Estos mensajes deben ser interpretados desde el punto de vista particular

del alumno, quien puede cometer errores debido a un conocimiento incompleto, o incluso incorrecto. Se espera que la interpretación de la retroalimentación supere estos errores cognitivos y promueva actualizaciones en el modelo de respuesta del alumno mediante la implementación de los procesos cognitivos y metacognitivos apropiados.

2. *Comportamiento adaptativo.* Nuestra teoría asume que el procesamiento de la retroalimentación es de naturaleza adaptativa. Los estudiantes deben evaluar la retroalimentación como un recurso relevante para alcanzar sus metas. Si consideran que esos mensajes contienen información relevante, entonces invertirán recursos cognitivos. Desde la perspectiva del aprendizaje autorregulado, los estudiantes deben dedicar su tiempo efectivamente para procesar la retroalimentación cuando sea necesario. Dado que responder preguntas de un texto es una actividad dirigida a un objetivo, bien comprender el texto, bien aprender unos conocimientos, los estudiantes han de monitorizar sus recursos y adaptarse al entorno para lograr sus objetivos.

3. *Recursos limitados.* Los recursos de procesamiento de los estudiantes son limitados. Por tanto, la forma en que se procesa la retroalimentación se puede explicar en parte como una función del análisis coste-beneficio. Los estudiantes deben evaluar si vale la pena participar en el proceso de retroalimentación, de modo que evalúan el coste cognitivo y emocional en relación con los beneficios que podrían obtener. Elaborar y monitorizar el uso de la retroalimentación requiere un procesamiento consciente que consume recursos de memoria de trabajo. Por tanto, los estudiantes deben estar relativamente motivados para procesar la EF a fin de lograr sus objetivos.

4. *Una cuestión de certeza en la respuesta.* Los estudiantes realizan juicios sobre la percepción de haber proporcionado la respuesta correcta. Estas evaluaciones internas pueden verse afectadas por la aparición de retroalimentación correctiva. La retroalimentación KR permite a los estudiantes verificar si sus evaluaciones eran correctas o no. Conocer si una respuesta es incorrecta parece desafiar los esquemas cognitivos del estudiante y, como consecuencia, desencadena la atención de los alumnos hacia la EF. Sin embargo, saber que la respuesta es correcta aumenta la certeza en la propia respuesta y permite verificar su conocimiento. En otros casos, un alumno puede dar la respuesta correcta con un nivel de confianza bajo, por lo que puede necesitar EF para corregir el conocimiento incompleto o erróneo adquirido. Por otro lado, cuando los estudiantes no reciben retroalimentación correctiva, el procesamiento de la EF se basa principalmente en las evaluaciones internas que realizan los estudiantes al completar una tarea.

5. Alcanzar el umbral. Relacionado con el supuesto anterior, asumimos que existe un umbral que desencadena los comportamientos de los alumnos como, por ejemplo, omitir la retroalimentación, prestar únicamente atención a la corrección de la respuesta o a la respuesta correcta, o buscar y procesar la EF cuando ésta es proporcionada. Cuando se alcanza ese umbral, el estudiante inicia un comportamiento u otro.

Además de estos supuestos, resulta igualmente importante prestar atención a los principios del “Marco de Componentes de Revisión del Conocimiento” (Knowledge Revision Components framework or KReC) propuestos por Kendeou y O’Brien (2014). Estos autores expusieron los procesos y mecanismos a partir de los que codificamos y reestructuramos nuestros conocimientos. En algunas ocasiones procesamos información que es congruente con nuestros conocimientos previos, por lo que esa información nos permite confirmar nuestro modelo mental de unos conceptos o ideas específicos. Sin embargo, en otras ocasiones nos enfrentamos a situaciones en las que hemos de codificar información que entra en conflicto con nuestros conocimientos. En esas situaciones, nuestros conocimientos pueden ser imprecisos o incluso erróneos, por lo que es necesario hacer ajustes en nuestras representaciones mentales. Para modificar ideas erróneas, el conocimiento erróneo y el correcto deben ser procesados conjuntamente para que puedan ser integrados y se produzca la revisión. Así, el marco KReC explica los mecanismos por los cuales las ideas correctas incrementan su activación y las incorrectas reducen su activación en la representación mental del estudiante. Estas situaciones, generalmente estudiadas mediante trabajos sobre refutación de conocimientos erróneos, pueden extrapolarse al procesamiento de la retroalimentación, momento en que nuestro modelo de respuesta puede ser confirmado o entrar en conflicto con la retroalimentación proporcionada. Si esta información es procesada activamente, entonces asumimos que los aprendices tal vez revisen y actualicen su modelo de respuesta de acuerdo con la información proporcionada en la retroalimentación. Este proceso de revisión es posible porque el modelo de respuesta inicial del estudiante y la información de la retroalimentación se activan en la memoria de trabajo al mismo tiempo, por lo que consideramos esencial proporcionar la retroalimentación inmediatamente después de responder cada pregunta. Por esta razón, consideramos que los principios propuestos por Kendeou y O’Brien (2014) pueden explicar el proceso de reestructuración de los conocimientos cuando se procesa la

retroalimentación formativa. A continuación se exponen brevemente los principios de este marco teórico y se ponen en relación con el procesamiento de la retroalimentación. Es importante señalar que estos procesos que aparecen a continuación pueden producirse siempre y cuando el estudiante haya tomado la decisión de participar activamente en su procesamiento.

1. Principio de codificación. Este principio asume que la información previamente aprendida puede interferir con el aprendizaje de nueva información, incluso cuando el aprendiz reconoce que la nueva información es correcta y entra en conflicto con su conocimiento. En el caso de la retroalimentación, la información proporcionada ha de ser codificada en el contexto de la tarea realizada. Así, puede que los conocimientos aprendidos por el estudiante entren parcial o totalmente en conflicto con la información de la retroalimentación.

2. Principio de activación pasiva. La información aprendida puede (re)activarse de forma pasiva, por lo que conocimientos relacionados con la información activa en la memoria de trabajo podrán activarse automáticamente. En el caso del procesamiento de la retroalimentación tras responder una pregunta, podemos asumir que el modelo de respuesta, así como otros conceptos del conocimiento previo relacionados con la tarea, estarán activos en la memoria de trabajo del estudiante cuando reciba la retroalimentación.

3. Principio de co-activación. Para que se pueda producir la revisión de conocimientos, es necesario que la información incorrecta y la información correcta estén activas en la memoria de trabajo al mismo tiempo. Al recibir la retroalimentación, podemos asumir que el modelo de respuesta y la información procesada de la retroalimentación estarán conjuntamente activas en la memoria de trabajo, por lo que podrá iniciarse el proceso de revisión de conocimientos.

4. Principio de integración. Una vez que se produce la co-activación y la nueva información entra en contacto con el conocimiento erróneo, entonces la nueva información correcta puede integrarse con la información errónea activa en la memoria de trabajo. Si esto se produce, la representación previa puede ser revisada en función de la nueva información procesada. En el procesamiento de la retroalimentación, la información correcta proporcionada puede integrarse con la información errónea del estudiante, siempre y cuando entren en contacto ambas informaciones en su memoria de trabajo.

5. Principio de competición de la activación. Si hay un incremento en la activación de la nueva información correcta, entonces ésta competirá con el conocimiento erróneo activo, disminuyendo su nivel de activación consecuentemente. Como apuntan los autores, incluso después del proceso de revisión del conocimiento, es posible que los estudiantes sigan reactivando información errónea previamente revisada que dé lugar a interferencias con nuevos aprendizajes. Una vez procesada la retroalimentación, esta información compite con el modelo de respuesta creado por el estudiante para responder la pregunta. Así, si la información correcta de la retroalimentación incrementa su activación y, al mismo tiempo, disminuye la activación del modelo de respuesta erróneo, la nueva información pasará a formar parte de la representación mental de los conocimientos aprendidos.

En definitiva, los supuestos y principios propuestos por ambas teorías, el modelo RESOLV y el marco KReC, nos permiten comprender las condiciones necesarias para que la retroalimentación formativa genere actualizaciones en el modelo de respuesta de los estudiantes cuando responden preguntas de un texto. Asumimos que estas actualizaciones generarán, a su vez, modificaciones en las representaciones mentales de los conocimientos evaluados con las preguntas. Seguidamente revisamos el papel de las diferencias individuales en la efectividad y procesamiento de la retroalimentación.

2.4. Diferencias individuales en la efectividad y el procesamiento de la retroalimentación

Una buena parte de la literatura reconoce el papel de las variables individuales relacionadas con la motivación, las habilidades lectoras, o el nivel de conocimiento previo, en el procesamiento y la efectividad de la retroalimentación (Hattie y Gan, 2011; Krause et al., 2009; Narciss, 2013; Narciss et al., 2014; Timms et al., 2016). Como sugirió Shute (2008) en su revisión, la literatura sobre retroalimentación muestra resultados inconsistentes que pueden explicarse por lagunas en la investigación en términos de interacciones entre las características de las tareas, los contextos en que son administradas y las características de los estudiantes (p. 176). De hecho, Goldin, Koedinger y Alevén (2012) encontraron que la efectividad de los mensajes de retroalimentación varía en función de los aprendices, por lo que concluyeron que debe existir una habilidad específica relacionada con el procesamiento de la

retroalimentación, y que ésta puede variar en función del nivel de complejidad de los mensajes. Por tanto, como se ha mencionado anteriormente, los estudiantes no sólo necesitan estar dispuestos a procesar los mensajes de retroalimentación (Aleven et al., 2003; Narciss, Koerndle, Reimann y Mueller, 2004), sino que además necesitan ser capaces de comprender esa información en el contexto de tarea asignado (Carless y Boud, 2018; Timmers y Veldkamp, 2011; Timms et al., 2016).

Recientes modelos han prestado atención a factores individuales que pueden intervenir en la efectividad de la retroalimentación formativa. Según Narciss (2013), factores individuales del aprendiz como, por ejemplo, la representación de estándares y requisitos de la tarea, sus habilidades de autoevaluación, sus estrategias de procesamiento de la información, o su disposición y habilidad para corregir errores, pueden intervenir en su efectividad. Respecto a las habilidades de autoevaluación, los estudiantes necesitan monitorizar su proceso de aprendizaje para generar retroalimentación interna y compararla con la retroalimentación externa recibida, por lo que la monitorización es un factor crucial en el procesamiento de la retroalimentación. En cuanto a las estrategias de procesamiento de la información, la habilidad del estudiante para procesar e integrar información de varias fuentes es un proceso esencial para poder establecer comparaciones entre el nivel alcanzado y el deseado. Por su parte, Timms et al. (2016) sugirieron que diferencias individuales en la motivación intrínseca, el conocimiento previo, la capacidad de la memoria de trabajo, la competencia lectora o la capacidad del lenguaje, pueden influir en cómo se procesa la retroalimentación formativa en entornos digitales. En otras palabras, la efectividad de la retroalimentación puede depender de la disposición y la capacidad de los estudiantes para usar la información ofrecida en los mensajes de retroalimentación. El modelo LP-FILE (Timms et al., 2016) tiene en cuenta las características individuales del alumno, de modo que las tres decisiones que los alumnos pueden tomar durante el proceso de retroalimentación podrían verse afectadas por diferencias individuales. Cuando se proporciona retroalimentación, el alumno puede continuar la tarea sin prestar atención al mensaje. Sin embargo, si el alumno decide leer la retroalimentación, los estudiantes con baja competencia lectora pueden tener dificultades para decodificar la información del mensaje. Finalmente, el conocimiento previo o la capacidad de la memoria de trabajo de los alumnos pueden influir en su capacidad para dar significado al contenido de la retroalimentación en relación con la tarea realizada.

La evidencia empírica previa muestra que el uso de la retroalimentación varía considerablemente entre los aprendices. Por ejemplo, Timmers y Veldkamp (2011) examinaron el uso opcional de mensajes EF en un entorno digital. Un grupo de estudiantes universitarios completaron una evaluación y recibieron un mensaje general indicando si sus respuestas eran correctas o incorrectas, i.e., retroalimentación KR. Desde esa pantalla con la corrección de sus respuestas, los estudiantes tenían la oportunidad de acceder a mensajes EF que contenían retroalimentación KCR y una explicación de los conceptos clave evaluados en cada pregunta. Los resultados mostraron que aproximadamente una cuarta parte de los estudiantes no accedieron a ningún mensaje adicional EF, y aproximadamente otro cuarto accedió a todos los mensajes EF en las preguntas respondidas incorrectamente. En un estudio posterior, Fox et al. (2014) investigaron los comportamientos de búsqueda de esta EF en relación con otras variables individuales, como el rendimiento en la prueba y la velocidad de trabajo en la tarea. En general, los estudiantes que obtuvieron un rendimiento mayor consultaron menos mensajes EF, principalmente si pertenecían a preguntas fáciles. Además, los investigadores observaron una relación positiva entre la atención dedicada a procesar la EF y la velocidad de trabajo, por lo que aquellos estudiantes que trabajan más rápido son los que procesan la EF más rápidamente.

La investigación existente sugiere que el conocimiento previo de los estudiantes, la competencia lectora, las habilidades metacognitivas, y la motivación pueden explicar los efectos de la retroalimentación (e.g., Fyfe, Rittle-Johnson y DeCaro, 2012; Gordijn y Nijhof, 2002; Krause et al., 2009; Maier et al., 2016; Narciss, 2008; Narciss et al., 2014; Schwartz, Tsang y Blair, 2016; Timmers et al., 2013). Por ejemplo, en el caso de la motivación, Maier et al. (2016) encontraron que la motivación intrínseca de los estudiantes predice el uso de la EF, lo que significa que los estudiantes con mayores niveles de motivación intrínseca dedican más esfuerzos a leer la retroalimentación y la valoran como una información relevante para la tarea. Igualmente, Timmers et al. (2013) encontraron que dos variables de tipo motivacional, i.e., las expectativas de éxito y el valor otorgado a la tarea, eran capaces de predecir la búsqueda de retroalimentación. A pesar de estos y otros hallazgos relacionados con factores motivacionales, los estudios que componen esta tesis tienen en cuenta la competencia lectora y el nivel de conocimiento previo de los estudiantes, ya que ambos factores determinan tanto la capacidad para construir una representación mental coherente como para integrar la nueva información en las estructuras de conocimiento previo.

Como señalamos en apartados anteriores, la competencia lectora de los estudiantes juega un papel esencial en la precisión del modelo de tarea construido al responder preguntas y en la habilidad para buscar información relevante (Cerdán et al., 2011; Gil et al., 2015; Vidal-Abarca et al., 2010). En el caso del procesamiento de la retroalimentación, existe escasa literatura que haya examinado el papel que juegan las habilidades lectoras (Golke et al., 2015), a pesar de que la mayoría de mensajes de retroalimentación proporcionados en entornos digitales suelen contener información escrita. En este caso, es posible que la competencia lectora explique la razón por la cual unos estudiantes son capaces de beneficiarse de esa información y otros no. En un estudio con aprendices de tercer grado de escuelas de preparación vocacional, Gordijn y Nijhof (2002) encontraron que los estudiantes con altas habilidades de comprensión se beneficiaron más de la EF que los estudiantes con bajas habilidades de comprensión. Schunk y Rice (1991, 1993) encontraron que estudiantes con dificultades de comprensión obtenían mejores niveles de comprensión lectora cuando recibían EF que les informaba sobre su progreso en el aprendizaje de estrategias para responder preguntas sobre los contenidos de textos expositivos (e.g., *Lo has comprendido bien porque has seguido los pasos en orden.*). Las diferencias entre estudiantes con alto y bajo nivel de comprensión han sido estudiadas en aprendices de inglés como segundo idioma. En este contexto, Bown (2017) encontró que los estudiantes con un nivel de lectura bajo obtuvieron un mejor rendimiento en la tarea tras recibir EF en comparación con aquellos que recibieron retroalimentación KR. Sin embargo, los estudiantes con un nivel de lectura alto obtuvieron un rendimiento similar independientemente de la retroalimentación recibida. Murphy (2007, 2010), sin embargo, no encontró diferencias entre aprendices de inglés con bajo y alto nivel de dominio de la lengua al completar una tarea de comprensión lectora en un entorno digital, independientemente del tipo de retroalimentación recibido: EF o KCR. El segundo estudio de esta tesis explora las operaciones mentales que realizan estudiantes de secundaria cuando procesan una EF específica sobre sus respuestas y la precisión para evaluar información relevante para la tarea, intentando localizar posibles diferencias entre estudiantes con alta y baja competencia lectora.

Por lo que respecta al conocimiento previo, ésta es, sin lugar a dudas, una de las principales variables a la hora de explicar el aprendizaje de los estudiantes, ya que representa la base sobre la que se construye el nuevo conocimiento (Alexander, 2003; Shapiro, 2004). Aunque existe escasa literatura que haya explorado la medida en que la

retroalimentación formativa influye en el aprendizaje de estudiantes jóvenes con diferentes niveles de conocimiento previo, varios estudios han demostrado que se trata de un factor que puede influir en su efectividad (Hannafin, Hannafin y Dalton, 1993; Smits et al., 2008). Por ejemplo, Smits et al. (2008) encontraron que los alumnos con niveles altos de conocimiento previo aprendían más con retroalimentación menos elaborada, aunque percibían más positivamente la retroalimentación más específica. Sin embargo, el nivel de especificidad de la retroalimentación no influyó en el rendimiento ni en las percepciones de utilidad de la misma en el caso de los alumnos con niveles bajos de conocimiento previo. En el área de las matemáticas, sin embargo, resultados recientes sugieren que estudiantes de primaria y universitarios con bajo conocimiento previo tienden a beneficiarse más de una intervención con retroalimentación que los participantes con alto conocimiento previo (e.g., Fyfe, 2016; Fyfe et al., 2012; Krause et al., 2009). Otros estudios, como el de Schwartz et al. (2016) han encontrado que los aprendices con bajo conocimiento previo pueden beneficiarse de recibir EF, mientras que los estudiantes con altos niveles de conocimiento pueden beneficiarse más tras recibir una retroalimentación correctiva más sencilla. En las situaciones de aprendizaje de conocimientos declarativos de ciencias que planteamos en los estudios 3 y 4 de esta tesis, los estudiantes cuentan con conocimientos erróneos que compiten con la información científicamente contrastada que aparece en los materiales (Chinn y Brewer, 1993). Cuando los estudiantes procesan la retroalimentación después de haber construido un modelo de respuesta total o parcialmente incorrecto, la información entra en conflicto con su conocimiento, y el resultado de esa competición puede producir cambios o actualizaciones en el modelo de respuesta inicial. Por tanto, el conocimiento previo de los estudiantes es un factor que puede determinar el uso y la efectividad de la retroalimentación formativa.

A pesar de la escasa investigación realizada sobre el papel de los factores individuales en el procesamiento y la efectividad de la retroalimentación formativa, los autores coinciden en señalar que estos factores pueden jugar un papel determinante, y reconocen la necesidad de examinar las relaciones entre las características de los estudiantes y la retroalimentación, especialmente cuando se trata de tareas que requieren procesos cognitivos de orden superior, como es el caso de la comprensión de textos (Golke et al., 2015).

3. USO Y PROCESAMIENTO DE LA RETROALIMENTACIÓN FORMATIVA PARA INCREMENTAR LA COMPRENSIÓN Y APRENDIZAJE DE CONOCIMIENTO CONCEPTUAL

A lo largo de esta introducción hemos revisado el papel de responder preguntas sobre los contenidos de los textos, así como las diferencias individuales más destacables al formar el modelo de tarea y la búsqueda de información relevante durante el proceso de resolución de las preguntas. Posteriormente hemos revisado el papel de la retroalimentación formativa en el aprendizaje de conocimientos conceptuales, especialmente cuando los mensajes son proporcionados en un medio digital. Asimismo, hemos planteado un marco teórico con el que comprender cómo se procesa la retroalimentación formativa que contiene información correctiva y elaborada en sus mensajes. Por último hemos examinado el papel de las diferencias individuales en el uso y la efectividad de la retroalimentación. A continuación presentamos el objetivo general y los objetivos específicos de los estudios de esta tesis.

3.1. Objetivo general

El objetivo general del presente trabajo consiste en analizar cómo interactúan los estudiantes de 3er grado de secundaria con unos mensajes de retroalimentación EF proporcionados por el sistema informático Read&Learn tras responder preguntas de comprensión y aprendizaje, así como sus efectos en el rendimiento, teniendo en cuenta la competencia lectora y el nivel de conocimiento previo de los estudiantes. Para ello realizamos cuatro estudios con objetivos específicos. Mientras que los dos primeros

estudios nos sirvieron para examinar el uso y procesamiento de una EF específica en una tarea que requería comprender ideas conceptuales de unos textos expositivos no académicos, los dos últimos estudios sirvieron para examinar la decisión de los estudiantes para acceder voluntariamente a unos mensajes EF con explicaciones sobre los contenidos de un texto expositivo de ciencias.

A partir de la revisión teórica planteada en la introducción de esta tesis, parece necesario estudiar en qué medida los estudiantes están dispuestos a procesar la retroalimentación formativa en tareas de responder preguntas de comprensión y aprendizaje en entornos digitales. Como hemos señalado, responder preguntas a partir de textos expositivos es una tarea que los profesores suelen asignar a sus alumnos muy frecuentemente para facilitar la comprensión y aprendizaje de sus contenidos (Ness, 2011; Sánchez y García, 2015; Sánchez et al., 2010) que implica tanto procesos de comprensión (e.g., realizar inferencias) como habilidades específicas en el manejo de textos (e.g., cómo buscar la información relevante para la tarea) (Anmarkrud et al., 2013; Cerdán et al., 2011; Gil et al., 2015; Mañá et al., 2017; McCrudden y Schraw, 2007; OECD, 2010; Roelle y Berthold, 2017; Rouet, 2006; Rouet y Britt, 2011; Rouet et al., 2017; Vidal-Abarca et al., 2010).

Al resolver las preguntas, los estudiantes suelen recibir retroalimentación, es decir, información sobre su rendimiento actual con respecto a un estándar para modificar sus pensamientos o comportamientos y así mejorar su aprendizaje (e.g., Hattie y Gan, 2011; Hattie y Timperley, 2007; Narciss, 2008; Shute, 2008). Estas tareas se ajustan a los entornos digitales, en los que docentes e investigadores pueden administrar tareas de comprensión y aprendizaje junto con retroalimentación formativa adaptada a la ejecución del estudiante en tiempo real (e.g., Azevedo y Bernard, 1995; Mason y Bruning, 2001; Mory, 2004; Shute y Rahimi, 2017). La efectividad de la retroalimentación ha sido estudiada en una amplia variedad de contextos y actividades de aprendizaje en entornos digitales, entre los que encontramos tareas de comprensión y aprendizaje de conocimientos conceptuales (e.g., Butler et al. 2013; Corbalan et al., 2010; Golke et al., 2015; Lee et al., 2009; Llorens et al., 2016; Maier et al., 2016; Moreno, 2004; Murphy, 2007). Atendiendo al estudio de la efectividad de la retroalimentación, hemos comprobado que la investigación previa muestra resultados inconsistentes en función del tipo de retroalimentación proporcionado, aunque, en términos generales, la EF es el tipo de retroalimentación más efectivo (e.g., Butler et al., 2013; Hattie y Timperley, 2007; Kluger y DeNisi, 1996; Mory, 2004; Narciss, 2004;

Shute, 2008; Van der Kleij et al., 2015). A pesar de la vasta investigación realizada sobre su efectividad, aún se desconocen los mecanismos relacionados con el procesamiento de la retroalimentación formativa en entornos digitales (Van der Kleij et al., 2012).

Los programas informáticos actuales son capaces de proporcionar mensajes EF tras responder cada pregunta, así como de registrar las conductas y tiempos de procesamiento durante su presentación. Habitualmente, los entornos digitales proporcionan retroalimentación automáticamente, esperando que el estudiante la procese también de forma automática. Sin embargo, su procesamiento requiere de un proceso activo de toma de decisiones que determina cómo los estudiantes hacen uso de la información proporcionada (e.g., Bangert-Drowns et al., 1991; Hattie y Timperley, 2007; Narciss, 2013; Timms et al., 2016). Así, los estudiantes han de monitorizar activamente su procesamiento mediante la búsqueda de correspondencia entre su nivel de comprensión y la retroalimentación para ajustar su representación al estándar proporcionado (e.g., Bangert-Drowns et al., 1991; Butler y Winne, 1995). Para ello necesitan implementar procesos cognitivos y metacognitivos (Carless y Boud, 2018), por lo que la retroalimentación puede tener un significado muy diferente en función de la interpretación que haga cada estudiante (Black y Wiliam, 2009). En este sentido consideramos que existen dos fases en el procesamiento de la EF que incluye algún tipo de retroalimentación correctiva. Inicialmente, los participantes tratan de verificar o refutar el modelo de respuesta generado durante la fase de resolución de la tarea. Posteriormente, los participantes deciden si necesitan revisar su conocimiento, por lo que aplican procesos de comprensión, monitorización y autorregulación que pueden producir actualizaciones en el modelo de respuesta validado. Por tanto, para que la EF produzca cambios en los conocimientos del estudiante, éste ha de estar dispuesto y ser capaz de comprender la retroalimentación recibida (Timmers y Veldkamp, 2011; Timms et al., 2016). Además, el uso y procesamiento de la EF puede estar influido por variables contextuales (e.g., presencia de retroalimentación correctiva o disponibilidad de los materiales) y características personales (e.g., competencia lectora o nivel de conocimiento previo). Por tanto, resulta esencial realizar investigación específica que analice la forma en que los estudiantes usan y procesan los mensajes EF en entornos digitales, al objeto de diseñar herramientas instruccionales más eficaces. Por ejemplo, los sistemas de tutorización inteligente como TuinLEC (Vidal-Abarca et al., 2014) o AutoTutor (Graesser et al., 2004), así como los nuevos libros de texto electrónicos,

pueden beneficiarse de los resultados obtenidos en el conjunto de estudios presentados en esta tesis.

3.2. Objetivos específicos

Para dar respuesta a este objetivo general, diseñamos cuatro estudios con objetivos específicos. Cada estudio se diseñó en función de los resultados encontrados en los estudios previos y en el corpus de investigación presentado en la introducción. A continuación se resumen los objetivos planteados en cada uno de los estudios. El objetivo del estudio 1 consiste en analizar el efecto de una EF específica que incluye información sobre el rendimiento al responder preguntas de comprensión y seleccionar información relevante (i.e., retroalimentación sobre qué información seleccionada es relevante e irrelevante y que, además, presenta el estándar con la información relevante al revisar el texto antes de pasar a la siguiente pregunta). Para ello pedimos a los participantes que respondieran un conjunto de preguntas de comprensión de unos textos expositivos no académicos y que seleccionaran la información textual que consideraban relevante para responder cada pregunta. Igualmente, analizamos en qué medida los estudiantes que reciben EF atienden a esos mensajes en comparación con un grupo control sin retroalimentación, controlando la influencia de la competencia lectora. El segundo estudio tiene como objetivo principal examinar cómo se procesa la EF administrada en el primer estudio, i.e., qué operaciones mentales aplican los estudiantes cuando reciben estos mensajes de retroalimentación. Los resultados de estos dos primeros estudios nos llevaron a realizar una serie de modificaciones en el diseño de los dos últimos estudios aquí presentados. El objetivo del estudio 3 consiste en conocer el efecto de la presencia de la retroalimentación correctiva KR y KCR en el acceso voluntario a mensajes EF que contienen explicaciones sobre los contenidos evaluados en las preguntas. En este caso empleamos un texto expositivo de ciencias y unas preguntas para fomentar el aprendizaje de sus contenidos, controlando el nivel de conocimiento previo de los estudiantes. Por último, el estudio 4 tiene como objetivo examinar el efecto de la retroalimentación KR y la disponibilidad del texto en el acceso voluntario a esos mensajes EF, ya que éstas son dos variables del contexto de tarea que pueden jugar un papel determinante en el uso de la EF. Asimismo, también examinamos el efecto del nivel de conocimientos previos y las decisiones de acceder a la EF en el rendimiento obtenido en las tareas administradas.

CHAPTER 2

STUDY 1: DOES COMPUTER-BASED ELABORATED FEEDBACK INFLUENCE THE STUDENTS' QUESTION-ANSWERING PROCESS?

Adapted from: Máñez, I., Vidal-Abarca, E., & Martínez, T. (2019). Does computer-based elaborated feedback influence the students' question-answering process? *Electronic Journal of Research in Educational Psychology*, 17(1), 81-106. doi:10.25115/ejrep.v17i47.2156

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 1° y 2° 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 control.

Discusión y Conclusión. 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

1. 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.

1.1. 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 (Cataldo & 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.

1.2. 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 allowed 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.

1.3. 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.

2. METHOD

2.1. 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.

2.2. 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

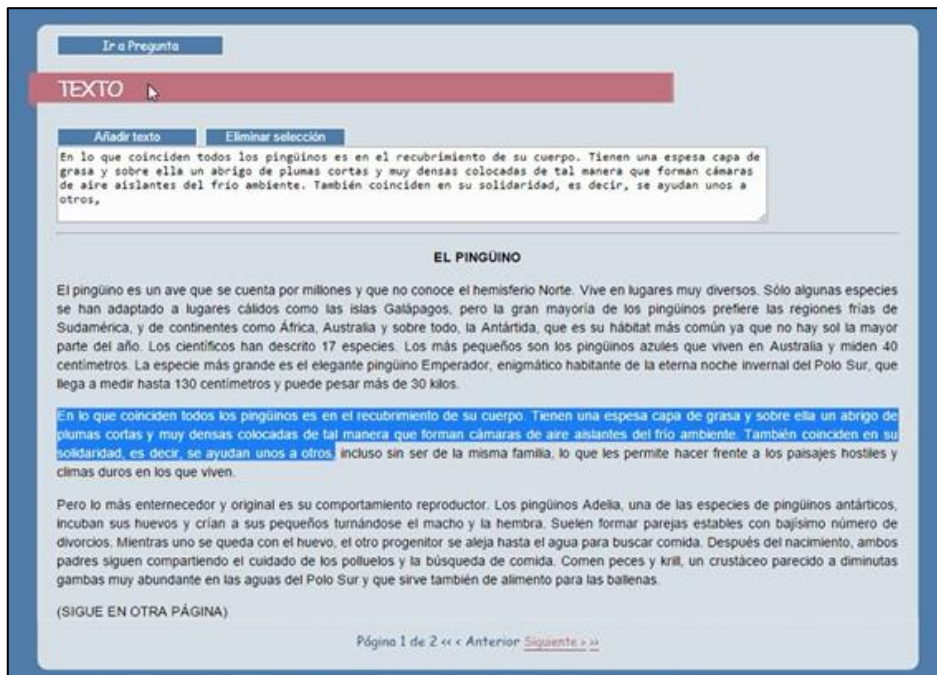
comprehension questions developed according to the PISA framework (OECD, 2006). Each correct response received a score of 1 and each incorrect a zero (maximum score: 20). The psychometric properties of CompLEC revealed acceptable reliability (*Cronbach's alpha* = .80) (Llorens et al., 2011).

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 according 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 information 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

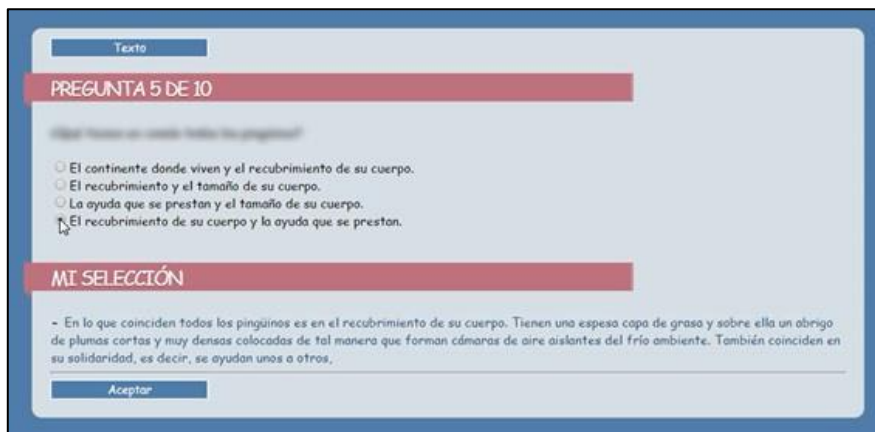


Figure 1. Screenshots of the task: (a) the text screen with a piece of text information selected, and (b) the question screen, with the question stem unmasked (above) and the answer choices unmasked (below).

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).

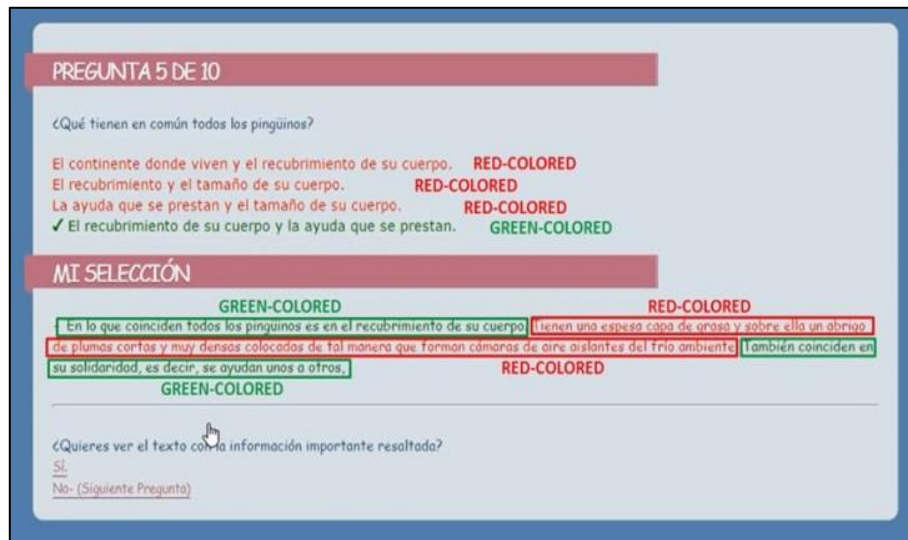
Once they closed that message, the question content with the students' selection was displayed on screen with visual feedback information. They received KCR on the answer to the question and KR on their selection, showing the correct choice in green and the wrong choices in red, as well as the relevant 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)

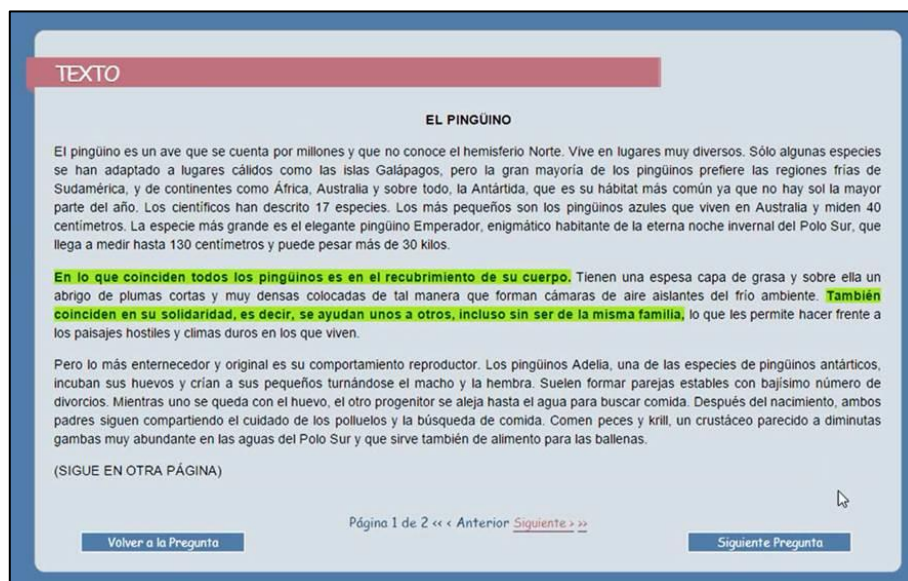


Figure 2. Screenshots of the EF messages: (a) pop-up box with written feedback, (b) KCR on the answer to the question and KR on the students' selection, and (c) text review screen with KCR on the standard question-relevant text information.

2.3. 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.

2.4. 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 computed 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.

3. RESULTS

3.1. 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, $F(1, 72) = 5.11, p = .027, \text{partial } \eta^2 = .07$, 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 information* selected, we found significant differences as a function of type of feedback, $F(1, 72) = 10.09, p = .002, \text{partial } \eta^2 = .12$, 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*, $F(1, 72) = 34.32, p < .001, \text{partial } \eta^2 = .32$, the *Question-relevant information* selected, $F(1, 72) = 31.14, p < .001, \text{partial } \eta^2 = .30$, and the *Non-relevant text information* selected, $F(1, 72) = 27.28, p < .001, \text{partial } \eta^2 = .28$.

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

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

3.2. 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).

	<i>Question Reading Time</i>	<i>Text-search Processing Time</i>	<i>Feedback Reading Time</i>	<i>Text Review Access</i>	<i>Text Review Reading Time</i>
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

4. 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 triggers 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., Papadopoulou, 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*. Paris: 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

CHAPTER 3

STUDY 2: HOW DO STUDENTS PROCESS COMPLEX FORMATIVE FEEDBACK IN QUESTION-ANSWERING TASKS? A THINK-ALLOUD STUDY

Adapted from: Máñez, I., Vidal-Abarca, E., Kendeou, P., & Martínez, T. (2019). How do students process complex formative feedback in question-answering tasks? A think-aloud study. *Metacognition and Learning*, *14*(1), 65-87. doi:10.1007/s11409-019-09192-w

ABSTRACT

The goal of this study was to examine how students process formative feedback that included corrective and elaborative information in online question-answering tasks. Skilled and less-skilled comprehenders in grade 8 read texts and answered comprehension questions. Prior to responding, students were asked to select the textual information relevant to answer each question. Students received formative feedback that included information about the answer correctness and an elaborative message, either automatically delivered or optionally accessed. Students' actions, as well as verbal protocols during feedback processing, were recorded. The results showed that: (a) students paid more attention to the answer correctness than to elaborative feedback messages; (b) students paid more attention to feedback information after failure than after success; (c) students actively monitored the accuracy of their responses; and (d) differences between skilled and less-skilled comprehenders in processing feedback were very limited. Theoretical and practical implications of these results are discussed.

Keywords: formative feedback; metacognitive processes; question-answering; reading comprehension; think-aloud.

1. INTRODUCTION

Question-answering supports learning during reading comprehension (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013; Cerdán, Gilabert, & Vidal-Abarca, 2011; Graesser & Murachver, 1985; Roelle & Berthold, 2017; Vidal-Abarca, Mañá, & Gil, 2010). This is an important reason why teachers often use question-answering activities in school settings (Ness, 2011; Pressley, Wharton-McDonald, Mistretta-Hampston, & Echevarria, 1998). In these contexts, teachers usually provide formative feedback about students' question-answering responses. Further, question-answering fits very well with computer-based instructional procedures for which formative feedback is a core component (e.g., Maier, Wolf, & Randler, 2016; Narciss et al., 2014). Although the effects of formative feedback have been extensively studied within computer-based learning environments (e.g., Hattie & Gan, 2011; Hattie & Timperley, 2007; Shute, 2008; van der Kleij, Eggen, Timmers, & Veldkamp, 2012), very little attention has been paid to understanding how students process this feedback. The present study aims to address this gap by examining how skilled and less-skilled comprehenders process formative feedback in question-answering tasks.

1.1. Feedback processing

Timms, DeVelle, and Lay (2016) have recently proposed a general decision-making model on how Learners Process Feedback in Intelligent Learning Environments (LP-FILE) that captures the steps in which students notice, decode, and make sense of feedback information. According to the LP-FILE model, an error has to be made initially in order for the student to take action. The student may detect and correct the error due to her "internal monitoring system". In the event that the student does not detect the error, the computer-based system tracks the error and delivers external feedback to the student, who has to notice and decode it, as well as to understand the feedback message in the context of the task performed. If the student processes this feedback appropriately, she would be able to correct her initial mistake. Therefore, according to the model, feedback is particularly sensitive to the student's mistakes because it is expected to reduce the gap between the current and the desired level of performance. Please note that similar processes to those just mentioned could be

triggered by the student's internal monitoring system when she detects an error before an external feedback message is provided.

However, LP-FILE is a general model and, as a result, it cannot adequately explain several of the nuances in feedback processing. First, LP-FILE does not specify what cognitive and metacognitive processes students activate when receiving computer-based formative feedback. In the present study, we explore the students' reactions to the feedback messages in an attempt to understand how students make sense of the feedback (*meaning making*), as well as how they *monitor* and *self-regulate* the use of this information. For example, do students *build meaning* by engaging in processes such as reading, paraphrasing or elaborating on the feedback information? Do they engage in *monitoring* by comparing their performance against the feedback? Do they engage in *self-regulation* by using additional information to meet the demands of the task? To examine these questions, we collected and coded verbal protocols to gauge students' cognitive and metacognitive processes deployed when they receive computer-based formative feedback. LP-FILE model does not acknowledge these questions and thus limits our understanding of how feedback may assist the student. Furthermore, although LP-FILE model acknowledges the potential effects of individual factors like reading competence or prior knowledge on feedback processing, it does not make concrete predictions about those effects.

Second, the LP-FILE model does not distinguish among different types of feedback (Shute, 2008) that impact performance and, consequently, may also impact processing. According to Shute (2008), Knowledge of Response (KR) feedback only indicates whether the students' answer is correct or incorrect, without making explicit what the right response is. This specific information is added when Knowledge of Correct Response (KCR) feedback is delivered (e.g., *the correct answer is X*). Elaborated Feedback (EF) includes information beyond the correct response, such as prompts (e.g., *please, think about X*), hints (e.g., *connect X to Y to give a response*), explanations (e.g., *the meaning of X is ...*), or worked-out examples (e.g., *an example of X is ...*), and it is usually delivered along with KR or KCR feedback (Narciss, 2008; Narciss & Huth, 2004; van der Kleij, Feskens, & Eggen, 2015). EF has been found to be more effective than KR and KCR feedback (e.g., Moreno, 2004; Mory, 2004; van der Kleij et al., 2015) because it allows learners to shorten the gap between the current and desired levels of performance. Finally, LP-FILE does not consider how feedback is processed during specific tasks. Previous research suggests that different tasks may

promote specific goals and processing that also has the potential to influence feedback processing (Bohn-Gettler & Kendeou, 2014). This is especially relevant in the context of the present study where we focus on feedback processing during question-answering tasks.

1.2. Feedback processing during question-answering tasks

Answering comprehension questions from an available text can be considered a hybrid task that includes both comprehension and problem-solving processes (Rouet, Britt, & Durik, 2017). Indeed, when students answer questions from an available text, they enact different strategies with the goal to provide the right answer to the question. Key components of these strategies are comprehension processes. A number of empirical studies provide data supporting these assertions. For example, Farr, Pritchard, and Smitten (1990) and Rupp, Ferne, and Choi (2006) used verbal protocols to examine how college students answered questions from an available text during completion of standardized reading comprehension tests (i.e., *Iowa Silent Reading Test* and *CanTEST*, in Farr et al., and Rupp et al., respectively). Both of these studies showed that participants directed resources (attentional focus and strategies) to identifying the best response to each question. To this end, participants used the questions to locate and evaluate answer-relevant information in the text. These findings are consistent with the self-regulation processes described by Vidal-Abarca, Mañá, and Gil (2010) for question-answering tasks. These processes include forming an appropriate task model for the question, making decisions about when to search the text, and in text-search processes.

Vidal-Abarca and colleagues conducted a series of studies to examine the role of feedback in text-search processes. In one study, Llorens, Vidal-Abarca, Cerdán, and Ávila (2015) found that elaborated feedback (EF) messages that included the question-relevant text information along with knowledge of response (KR) feedback improved the students' performance on inference questions and self-regulated text-search strategies. In a follow-up study, Llorens, Vidal-Abarca, and Cerdán (2016) examined the effects of two different types of elaborative feedback. In one condition, students were asked to select the text information they considered relevant to answering the question prior to responding, whereas in the other condition students answered the questions without having to select relevant text. After answering each question, students

in both conditions received knowledge of response (KR) and knowledge of correct response (KCR) feedback, as well as elaborative feedback (EF) about their text-search processing (i.e., relevant information read or selected, respectively). Additionally, students in the text selection condition had the option to access the ‘ideal question-relevant text’ selection so that they could compare their selection with the ideal selection. Llorens et al. (2016) found that students in this condition not only benefited immediately from the feedback they received, but also demonstrated superior performance after a 2-h delay on a transfer task. Llorens et al. showed that it was the feedback associated with the text selection task that produced these benefits, not the selection task itself.

Additionally, Máñez, Vidal-Abarca, and Martínez (2016) found that providing students with EF information improves not only their question-answering performance, but also their efficiency in evaluating textual relevance compared to a control group with no feedback. Researchers found that both groups devoted the same amount of time to searching the text for question-relevant text information, although the students who received EF selected less non-relevant text information than the students in the control group. These findings are pertinent to the present investigation as we adopted the EF condition mentioned above (Máñez et al., 2016; Llorens et al., 2016) in an effort to uncover the exact meaning making, monitoring, and self-regulation processes at play when students engage with this kind of feedback.

Students’ decisions to access feedback can be very informative regarding how feedback functions in digital learning environments. For example, Timmers and Veldkamp (2011) and Fox, Klein Entink, and Timmers (2014) asked students to answer a set of questions on information literacy. Then, the computer-based system delivered an overview of knowledge of response (KR) feedback for each answer. Afterwards, participants had the opportunity to access additional elaborative feedback (EF) for each question that included knowledge of correct response (KCR) and an explanation. Findings indicated that students’ attention to EF varied as a function of answer correctness. Specifically, students mainly accessed EF for incorrectly answered questions, suggesting that students tend to use feedback for its corrective function.

In a similar vein, Maier et al. (2016) evaluated the effects of feedback type on performance within a computer-assisted formative assessment. They had secondary-school students study and answer questions about a biology text under three feedback conditions: EF (i.e., KR feedback and a short explanation after incorrectly answered

questions), KR feedback, and no feedback. Students in the EF condition also self-reported their use of EF. Initial results showed that there were no significant differences in student performance between conditions. However, when researchers divided the students in the EF condition into those who self-reported to have used and have not used the feedback, they found that students who received and used EF, as well as those who received KR feedback, outperformed students in the no-feedback condition and students who received EF but did not use it. This finding suggests that the active use of EF plays a key role during learning and raises questions about individual differences in feedback use and processing.

1.3. Individual differences

As Shute (2008) posits in her review, “inconsistent findings with regard to feedback effects on learning may be explained by research gaps in terms of interactions among task features, instructional contexts, and student characteristics” (p. 176). In fact, the LP-FILE model (Timms et al., 2016) acknowledges that individual learner characteristics, such as reading competence, prior knowledge, intrinsic motivation, or working memory, may influence the decisions learners make when feedback is delivered (i.e., noticing, decoding, and making sense). Students may decide not to pay attention to feedback, but if they do, those with poor comprehension skills may struggle to decode its content. Maybe these students would make fewer decisions to access the feedback to begin with. Examining individual student differences in the context of feedback processing is an emerging area of research. Fox et al. (2014) investigated students’ feedback-seeking behaviors in relation to individual variables such as test performance and processing speed. Findings showed that students who obtained higher performance scores consulted fewer elaborative feedback messages that pertained to easy questions. Additionally, there was a positive relation between students’ attention devoted to elaborative feedback and processing speed, indicating that the faster students work, the faster they process feedback.

An important individual difference that has the potential to influence feedback access and processing is comprehension skill. Indeed, Vidal-Abarca, Mañá, and Gil (2010) showed that skilled-comprehenders are more efficient and effective in forming an appropriate task model for the question, making decisions about when to search the text, and engaging in text-search processes during question-answering tasks.

Furthermore, the text and discourse literature has consistently highlighted text processing differences between skilled and less-skilled comprehenders (Kendeou, McMaster, & Christ, 2016; Perfetti, 1985). For instance, previous findings suggest that skilled comprehenders are more accurate in their understanding of the task demands (Cerdán, Gilabert, & Vidal-Abarca, 2013), as well as in locating question-relevant information when answering questions from an available text (Cerdán et al., 2011). These findings, taken together, suggest that comprehension skill is an important individual difference to consider when examining feedback processing.

1.4. The current study

The goal of this study is to examine how students process formative feedback that includes both corrective and elaborative information during question-answering tasks, and to explore potential individual differences between skilled and less-skilled comprehenders in feedback processing. Specifically, using a think-aloud methodology (Krause, Stark, & Mandl, 2009) and a computer-based environment (Read&Learn software; see Methods for a detailed description) we examined moment-by-moment how students processed complex feedback messages when answering comprehension questions for two texts. We used a task that required students to select the text information relevant to answering each question prior to providing an answer (Máñez et al., 2016; Llorens et al., 2016). Feedback messages included knowledge of response (KR; correct vs. incorrect) and knowledge of correct response (KCR; correct answer), as well as elaborative feedback (EF) messages. EF messages were either automatically delivered (i.e., a summary of the student's text-relevant selection accuracy, plus a prompt for how to proceed on the following questions) or optionally accessed (i.e., a view of the ideal text-relevant selection). The whole task was performed in a computer-based environment that recorded students' actions while processing feedback and thinking-aloud (in one of the two texts). The students' online actions and their verbalizations were recorded and transformed into behavioral and verbal data indicative of cognitive and metacognitive processing (Dunlosky & Metcalfe, 2009).

Drawing on the extant literature (Llorens et al., 2016; Rouet et al., 2017; Timms et al., 2016), we made several predictions. First, we predicted that students would focus their attention more on answer correctness (i.e., KR and KCR) than elaborative feedback (EF) messages. Even though it may be reasonable to expect that skilled

comprehenders' strategies may also influence feedback processing, we did not have an a priori prediction about *how* skilled and less-skilled comprehenders would differ in their focus of attention on feedback. Second, we predicted that students would focus more attention on elaborative feedback messages after failure than after success. We anticipated that there would be differences between skilled and less-skilled comprehenders in this regard because less-skilled comprehenders would need to spend more time to understand the feedback information compared to their skilled counterparts. Third, we predicted that students' meaning making cognitive processes (e.g., paraphrases and elaborations) and metacognitive processes (e.g., comparing their current response to the correct answer, assessing their comprehension, or self-regulating their actions) would be positively related to comprehension performance. Regarding individual differences in the cognitive and metacognitive operations, we expected that skilled comprehenders would deploy more meaning making operations than their less-skilled counterparts because the skilled ones are able to better understand the task demands while answering comprehension questions from an available text (Cerdán et al., 2013). Conversely, we expected that less-skilled comprehenders would deploy more monitoring reports than their skilled counterparts because they may struggle to understand the feedback information within the task performed.

2. METHOD

2.1. Participants and design

A total of fifty-two 8th-grade students with a mean age of 13.38 years ($SD = .91$) from two public schools in Spain participated in this study (48% were female).

The study used a 2 (Comprehension skill: skilled vs. less-skilled) X 2 (Thinking-aloud: thinking-aloud vs. no thinking-aloud) mixed design with comprehension skill as a between-participants variable and thinking-aloud as a within-participants variable. Comprehension skill was assessed with CompLEC (Llorens et al., 2011), a standardized test that included five texts and twenty comprehension questions. Each correct response received a score of 1 and incorrect responses a zero (maximum score: 20). Based on these scores, we performed a tripartite split and selected 20 skilled ($M = 16.10$; $SD = 1.45$) and 20 less-skilled ($M = 8.55$; $SD = 1.96$) comprehenders to maximize differences between groups, $t(38) = 13.86$; $p < .001$. The data of eleven students with intermediate

reading comprehension scores and one student who received special education services were excluded from subsequent analyses.

Participants read two texts and answered 10 multiple-choice comprehension questions per text. We counterbalanced the order of the texts and the order of thinking-aloud to avoid any interference between these two factors. No order effects of text or think-aloud were found on CompLEC ($F(1, 36) = .001; p > .05; F(1, 36) = .01; p > .05$, respectively) or task performance ($F(1, 36) = .28; p > .05; F(1, 36) = .28; p > .05$, respectively). Participants performed the question-answering task individually in a quiet room in their school. Students' online behaviors and thinking-aloud verbalizations were recorded during processing of the feedback messages.

2.2. Materials

2.2.1. Texts and questions

Materials included two expository texts with ten multiple-choice comprehension questions per text from the standardized Test of Comprehension Processes (TCP; Martínez, Vidal-Abarca, Sellés, & Gilabert, 2008). The two texts were 550 and 471 words long. Flesch-Szigriszt readability level was 66.89 for one text and 63.39 for the other, indicating that they can be considered "Quite easy" and "Normal" to read according to the INFLESZ scale. Texts' coherence and organization were analyzed using ETAT (Expository Text Analysis Tool; Vidal-Abarca, Gilabert, & Abad, 2002a; Vidal-Abarca et al., 2002b), which allowed researchers to determine which sections of the text could be considered relevant to answering each question. The connection 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, Vidal-Abarca, Gil, & Gilabert, 2009). Text segments of relevant information to answering each question 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 score of zero (maximum score: 20). The psychometric properties of the test revealed moderate indices of reliability (*Cronbach's alpha* = .72) and validity (*Pearson's r* of the computerized version of the test and scores on the comprehension subtest of PROLEC-SE = .64).

Sample Question – Text Selection – Feedback Types (see Figure 2 for the Apparatus section). For the question 'What do all the penguins have in common?' the

correct response was *'The coating of their bodies and the help they provide'*. The relevant text information included the following sentences: (1) *All penguins share the covering of their bodies.* (2) *They are also supportive, that is, they help each other, even if they are not from the same family.* Thus, in order to provide the correct response, the student needed to integrate information from these sentences. Based on students' own selection of relevant information and answer, several feedback messages were automatically delivered. If a student selected too much information (e.g., selected additional unnecessary information from this or other paragraphs) and provided an incorrect response, then the computer would deliver the following feedback: *Incorrect. You have selected more information than what was necessary. Review the alternatives. Also, review in the text what pieces of unnecessary information you have selected. Next time, avoid selecting unnecessary information.* If a student selected some relevant information and provided the correct response, then the computer would deliver the following feedback: *Correct. You have selected some relevant information. Review in the text what information you missed. Next time, make sure you select all the necessary information.* These messages were designed this way because the odds of providing an incorrect response after selecting one or several pieces of non-relevant information are higher than after missing a piece of relevant information. Including non-relevant information indicates that the student assesses as relevant a piece of information that has little or nothing to do with the question, which may suggest that the student is not on the right track. However, discarding some question-relevant information may suggest that the student is on the right track and able to provide an answer with less information available. The first part of each of these feedback messages corresponds to knowledge of response (KR) feedback (correct vs. incorrect), whereas the second part is an elaborative feedback (EF) message that includes: (a) a summary of the current selection of information, (b) a hint on how to proceed when reviewing the answer and the selection of relevant information for this specific question, and (c) an advice on how to search the text when responding to the questions.

2.2.2. Apparatus

Participants completed the reading comprehension tasks (i.e., selecting relevant text information and answering questions) and received feedback on Read&Learn, which is the updated version of Read&Answer (Vidal-Abarca et al., 2011), a software program that records online behaviors and processing times during the completion of

reading comprehension tasks. Read&Learn provides students with different feedback messages according to the students' online behaviors and performance.

Texts and questions were presented on different screens (see Figure 1). Each text was presented in two different pages because of its length. At the top left corner of the question screen, the "Text" button allowed students to return to the text screen while answering the questions (Figure 1a). At the top left corner of the text screen, the "Question" button allowed students to return to the question screen (Figure 1b). This interface provided flexibility to the question-answering process since students had to assess and select the pieces of text information germane to each question. Multiple-choice question content was displayed using a masking procedure. Once the question wording was unmasked, the choices remained masked, and vice versa (Figure 1a). The masking procedure was also used by Máñez et al. (2016) and Llorens et al. (2016) and served to ensure that students processed all question components. After reading the questions, students went to the text screen to highlight (using the computer mouse) the information they considered relevant for answering the question (Figure 1b). The system allowed students to select words, sentences, or entire paragraphs. Students could go back and forth between the question and the text screens.

a) Question screen



b) Text screen

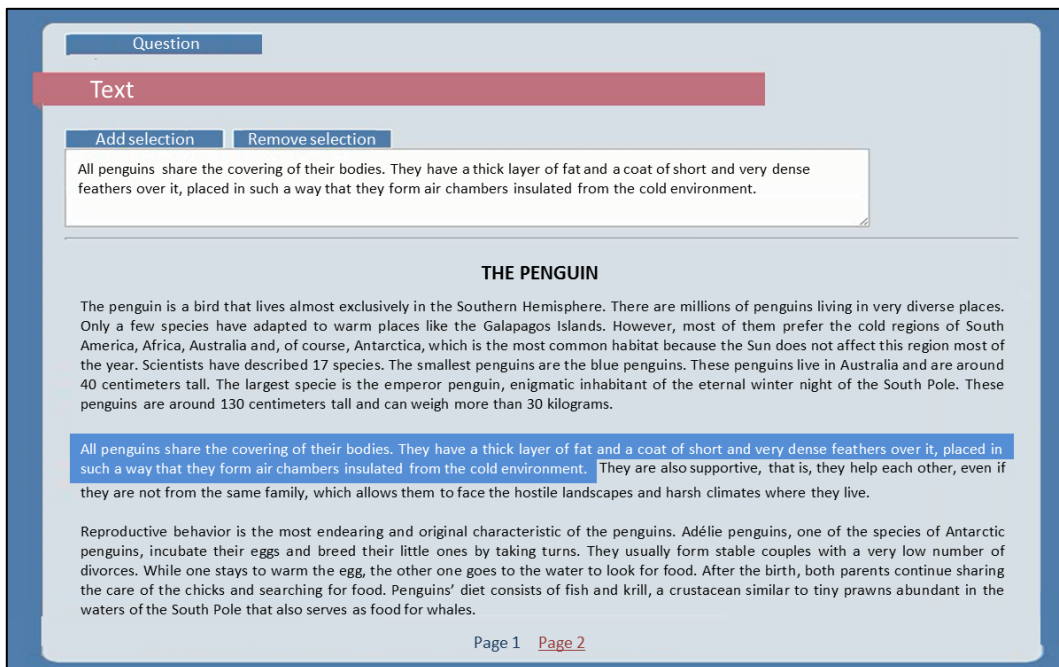
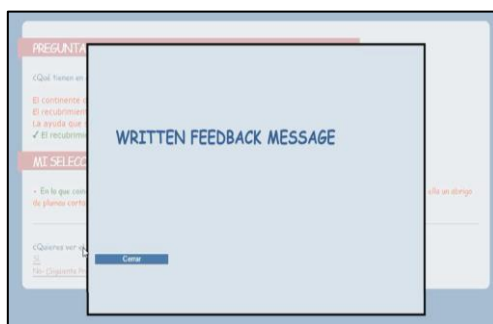


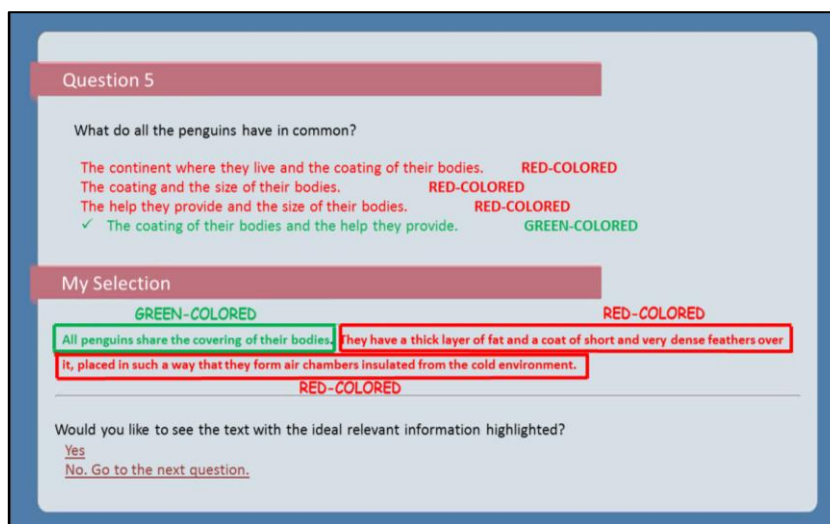
Figure 1. Screenshots of the multiple-choice question screen with the question stem unmasked (above) and the answer choices unmasked (below) (a), and the text screen with a piece of text information selected (b). Please note that the displayed information in My selection section (a) is the same as the information selected in the text.

After answering each question, Read&Learn provided students with feedback messages according to their performance. Read&Learn first delivered a written feedback message in a pop-up box consisting of the knowledge of response (KR) feedback about the correctness of the response and the elaborative feedback (EF) about the relevant information selected (Figure 2a). Once students closed the pop-up box, they could see knowledge of correct response (KCR) feedback in the upper part of the screen (the correct response in green and the incorrect responses in red), and knowledge of response (KR) about the current selection of information (with the relevant information depicted in green and the non-relevant information selected in red; Figure 2b) in the lower part of the screen. Following these automatically-delivered feedback messages, students had the option to access an elaborative feedback (EF) message on the text review screen with the ideal question-relevant text information highlighted (Figure 2c). Students were free to move from 2c to 2b screen and vice versa to directly compare their text selection to the ideal text selection. Students could move to the next question by clicking on the “Next Question” button.

a)



b)



c)

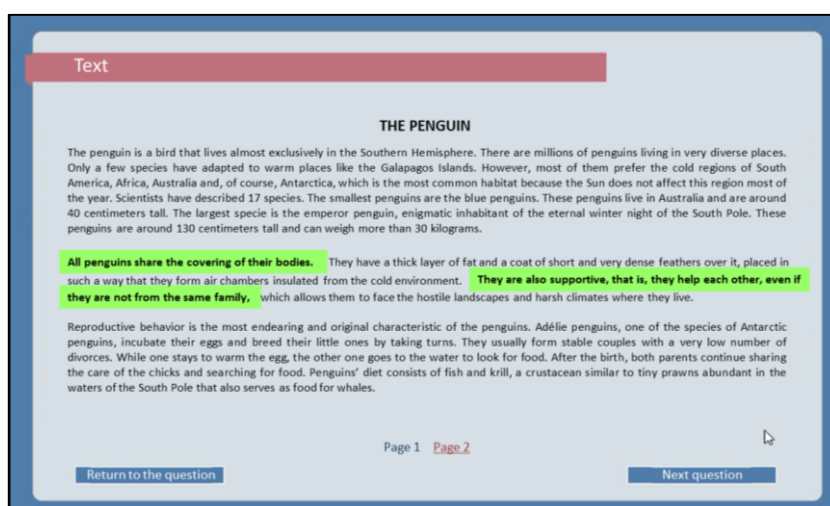


Figure 2. Screenshots of the feedback messages: (a) pop-up box with KR plus EF feedback messages (e.g., *Incorrect. You have selected only part of the necessary information. Review the text to see what information you missed. Next time try to select all the relevant information*), (b) KCR over the question answer and details of the current selection of information mentioned in the previous EF message, and (c) screenshot of the EF feedback message consisting on the text review screen with the ideal question-relevant text information.

We used CamStudio (www.camstudio.org) to record students' thinking-aloud verbalizations when feedback messages were delivered. This program records both the participants' voices and the personal computer screen.

2.3. Procedure

The experiment tasks involved two sessions of approximately one hour each. In the first session, students' comprehension skills were assessed using a paper-and-pencil version of the CompLEC test; students completed the test in their regular classroom. In the second session, each student was tested individually in a quiet room in his or her school. Students were informed that they would have to read the text before answering the questions and return to the text screen to perform the selection task while answering each question. Students read the first text. Then, they answered the corresponding 10 questions receiving formative feedback after each question. Next, students repeated the procedure with the second text and the corresponding 10 questions. For each question, students had to select the pieces of text information they considered relevant to providing an answer. Students completed the task for one text silently, whereas they thought-aloud for the other text (counterbalanced order).

Task completion was supervised by a trained experiment leader who instructed students on how to use the Read&Learn software program and how to report their thoughts. The following instructions were presented orally: *“You are going to read two texts and answer 10 questions for each. You will have to select the text information you consider relevant to answering each question. After performing both of these tasks, the program will automatically provide you with feedback on your performance. Further, you will have the opportunity to access the text screen with additional feedback on the correct question-relevant text information, if you wish. You will perform the task in silence in one text, but you will have to tell me everything that goes through your mind when processing the feedback in the other text. If you do not say aloud what you are thinking about for 5 s, I will remind you to do so. Now, I am going to show you how the program works and how to report your thoughts using a couple of examples (we showed participants a brief text and two questions performed by a fictitious student who reported meaning making, monitoring and self-regulation reports). Remember, this is just an example. There are no correct or incorrect thoughts. It is very important that*

you say aloud everything that goes through your mind. Do you have any questions? Thank you for participating". Thus, the experiment leader emphasized that verbalizations were just examples and encouraged the students to report everything that came to minds. If a student did not think-aloud during the task for a period of 5 s, the experiment leader asked "*What are you thinking right now?*" Following standard think-aloud procedures (Magliano & Millis, 2003), the experiment leader did not provide feedback regarding the content of the students' thoughts. Students' verbalizations were recorded and subsequently transcribed for protocol analysis. Only the responses of the students who were selected in the two comprehension skill groups were coded.

2.4. Processing measures

Processing measures were extracted from (a) the think-aloud protocols and (b) the Read&Learn log files of students' sequence of actions.

2.4.1. Think-aloud protocols

Students' responses during the think-aloud procedure were transcribed and coded following Cerdán and Vidal-Abarca (2008). Think-aloud comments could reflect one or more kinds of thoughts (Wolfe & Goldman, 2005). Thus, students' verbalizations were parsed into idea units. Idea units constitute one or more clauses containing a meaningful verbalization that reflects meaning independently of other idea units in the comment. Idea units varied in length. Some utterances were relatively short (e.g., "*I was right*") and others were more elaborated (e.g., "*I do not remember why the penguins help other penguins, so I am going to review the text and see where this information is*").

Each idea unit was categorized on the basis of a coding scheme adapted from previous research (e.g., Ramos & Vidal-Abarca, 2013). The coding scheme identified the mental operations students engaged in during feedback processing. It consisted of the following four main categories: meaning making, monitoring, self-regulation, and others. *Meaning making* involves comprehension processes that help build understanding (i.e., restatements, paraphrases, and elaborations). *Monitoring* involves evaluation of a student's own understanding (e.g., comparing the student's performance against the external feedback, affective reactions, assessing one's comprehension). Monitoring may alter the student's knowledge and guide self-regulation. *Self-regulation* involves actions that a student engages in to meet the demands of a task (e.g., seeking

external feedback, asking for explanations). *Others* is a category that includes any response that did not fall into any of the categories described above. As can be seen in Table 1, meaning making and monitoring categories were divided into subcategories that included more detailed information (see examples of each category in parenthesis). This coding scheme allowed us to evaluate the cognitive and metacognitive processes by computing the proportion of verbalizations each student reported for each category. Thus, the sum of proportions across these categories (i.e., meaning making, monitoring, self-regulation, and others) does add up to 1.

Table 1. *Think-aloud categories for processing the formative feedback messages*

a) Meaning making: Students' verbalizations aimed at building knowledge.

Restatement: (re)readings of units of information.

Paraphrasing: transformations of units of information involving either order word changes, synonyms, omissions or repetition of words (*"I have selected some of the relevant information"*, *"It was because they had not been invaded"*).

Elaboration: interpretations of information adding meaning to the explicit content (i.e., making inferences, giving examples, or explaining causal relationships) (*"So both of them incubate the egg and not just one"*, *"It is a prawn that feeds only the penguins"*, *"The penguins help the small, the adults help the small, while parents go to look for food"*).

b) Monitoring: Students' evaluations of their own understanding.

Performance comparison: verbalizations about the relationship between students' accuracy and a standard (*"It was the other response with which I was hesitating"*, *"So, I selected more than I needed"*, *"Almost the entire paragraph has question-relevant information and I did not select it"*).

Affective reaction: verbalizations about students' perceptions regarding how well or how poorly they are performing the tasks and verbalizations about liking or disliking it (*"Perfect!"*, *"I am doing it well so far"*, *"Today is not my day"*, *"This is annoying"*).

Attribution: verbalizations about possible explanations regarding the cause of student's performance (*"I was wrong because I have not read well"*, *"Okay, so I misunderstood the question, I guess"*, *"I was right because it was very easy"*, *"I failed because it is very poorly explained"*).

Assessing comprehension and memory: reading comprehension, lack of knowledge and memory assessments (*"I get it"*; *"It makes sense"*; *"I remember now"*, *"I do not understand that"*; *"I do not know what it means"*).

c) Self-regulation: Verbalizations about actions that a student engages in to meet the demands of a task (e.g., seeking external feedback, asking for explanations). For example, students' verbalizations about actions to solve a specific problem previously stated (*"Let's see what information I selected that I should not have selected"*; *"I am going to see what I forgot to select"*, *"I am going to look over the text because I do not understand why I failed"*).

d) Others: Students' meaningful reports that did not fall into any of the categories described above, reports with an unclear meaning, or irrelevant reports such as software usage or mandatory behaviors (*"I thought this one was correct"*, *"I did not know it was here"*, *"They should not have killed the Sioux"*, *"Nothing comes to my mind"*, *"Moving on to the next question"*).

Additionally, we coded where students focused their attention on when processing feedback information. This coding (*feedback focus of attention*) allowed us to evaluate which parts of feedback students attended to by computing the proportion of questions in which students verbalized about: (a) the correctness of the answer choices, corresponding to KR or KCR feedback, (b) the accuracy of the student text selection, corresponding to EF delivered automatically, and (c) the ideal question-relevant text information, corresponding to the optional EF. Each of these categories is independent from the others and a student might focus on more than one feedback component. For example, a student may verbalize about the correctness of the answer, as well as about the accuracy of the text selection or the ideal text selection in a specific question. Thus, the sum of proportions across these three categories (i.e., Answer correctness, Current text selection, and Ideal text selection) does *not* add up to 1.

Following Cerdán and Vidal-Abarca (2008), two coders initially trained together. Then, they independently parsed and coded 10% of the protocols. Cohen's kappa for parsing was perfect ($K = 1$; $p < .001$) and acceptable for coding ($K = .72$; $p < .001$). Disagreements were resolved through discussion, and the remaining protocols were parsed and coded by the first coder.

2.4.2. *Read&Learn - Behavioral measures*

We obtained the following online measures from Read&Learn to better understand the students' focus of attention: (a) *Feedback reading time* was the total time students devoted to processing the automatically-delivered feedback messages (i.e., first two feedback screens); (b) *Text review access* was the proportion of questions in which students decided to access the text review screen with the ideal question-relevant text information highlighted, and (c) *Text review reading time* was the total time students spent processing the optionally-delivered feedback with the ideal question-relevant text information. All of these measures were computed conditionally to question-answering success (i.e., correct or incorrect responses) and to the think-aloud condition (i.e., thinking-aloud or no thinking-aloud) in order to provide a richer understanding of the attention paid to processing the feedback.

3. RESULTS

The results are organized by the three main predictions. First, we report on the students' focus of attention on feedback. Second, we report on attention paid to feedback depending on the question-answering failure or success. Finally, we report on the cognitive and metacognitive processes deployed when processing feedback and the relation between online feedback processing and students' performance.

3.1. Focus of attention on feedback information

We conducted a mixed ANOVA with *feedback focus of attention* components (answer correctness, current text selection, ideal text selection) as a within-subjects variable, and comprehension skill (less-skilled, skilled) as a between-subjects variable. Table 2 reports the means and standard deviations.

Table 2. *Descriptive Statistics of the Proportions of the Feedback Focus of Attention, as a Function of Comprehension Skill (Less-skilled, Skilled)*

Comp. Skill	Answer Correctness	Current Text Selection	Ideal Text Selection
Less-skilled	.69 (.22)	.41 (.27)	.32 (.30)
Skilled	.63 (.29)	.39 (.28)	.30 (.21)

Note. *SD* in parentheses.

There was a main effect of the feedback focus of attention with respect to the amount of reports that referred to the three feedback type components, $F(1, 38) = 35.66$; $p < .001$; $\eta^2 = .39$. Pairwise comparisons revealed that participants paid more attention to the correctness of the answers ($M = .66$; $SD = .25$) than to the accuracy of their current selection ($M = .40$; $SD = .27$), $t(38) = 4.48$; $p < .001$, or the ideal text selection ($M = .31$; $SD = .25$), $t(38) = 6.00$; $p < .001$. Moreover, students verbalized more often about the accuracy of their current selection than the ideal text selection, $t(38) = 2.44$; $p = .02$. There was no interaction between the students' feedback focus of attention and comprehension skill level, $F(1, 38) = .09$; $p > .05$ nor a main effect of comprehension skill, $F(1, 38) = .37$; $p > .05$.

These results were consistent with our predictions. Students, independent of their comprehension skill, were mainly interested in knowing the answer correctness rather than getting information about the accuracy of their current text selection.

3.2. Attention on feedback conditional to question-answer (success vs. failure)

We conducted three mixed ANOVAs with comprehension skill (less-skilled vs. skilled) as a between-subjects variable, and question-answering (success vs. failure) and think-aloud method (thinking-aloud vs. no thinking-aloud) as within-subjects variables for each behavioral feedback measure (i.e., *Feedback reading time*, *Text review access* and *Text review reading time*). Table 3 reports the means and standard deviations for each of these measures.

Table 3. *Descriptive Statistics of the Feedback Reading Time, Text Review Accesses, and Text Review Reading Time as a Function of Question-Answer (Success, Failure), Think-aloud Method (Think-aloud, No Think-aloud), and Comprehension Skill (Less-skilled, Skilled)*

		Success		Failure	
		Think-aloud	No Think-aloud	Think-aloud	No Think-aloud
	Comp. Skill				
Feedback Reading Time (in seconds)	Less-skilled	10.54 (4.46)	6.67 (2.56)	18.71 (8.02)	13.31 (6.75)
	Skilled	8.39 (3.44)	5.54 (1.64)	16.19 (7.22)	10.31 (4.55)
Text Review Access (proportion)	Less-skilled	.36 (.39)	.31 (.32)	.74 (.28)	.67 (.31)
	Skilled	.35 (.36)	.30 (.30)	.77 (.34)	.91 (.15)
Text Review Reading Time (in seconds)	Less-skilled	11.33 (7.66)	5.75 (2.65)	13.00 (6.66)	7.15 (2.57)
	Skilled	12.14 (9.25)	7.32 (4.80)	12.56 (5.67)	11.03 (5.93)

Note. *SD* in parentheses.

For the *Feedback reading time* (in seconds), there was a main effect of question-answering success, $F(1, 35) = 68.49$; $p < .001$; $\eta^2 = .34$. Pairwise comparisons revealed that participants spent more time processing the feedback after failure ($M = 14.67$; $SD = 5.28$) than after success ($M = 7.81$; $SD = 2.72$), $t(35) = 8.28$; $p < .001$. Moreover, there was a main effect of think-aloud method, $F(1, 35) = 31.15$; $p < .001$; $\eta^2 = .15$. Pairwise comparisons revealed that participants spent more time processing feedback when

thinking-aloud ($M = 13.49$; $SD = 4.56$) than when not thinking-aloud ($M = 8.99$; $SD = 3.71$), $t(35) = 5.58$; $p < .001$. Further, there was a significant main effect for comprehension skill, $F(1, 35) = 4.28$; $p < .05$; $\eta^2 = .11$, with less-skilled comprehenders ($M = 12.31$; $SD = 3.69$) spending more time processing feedback than their skilled counterparts ($M = 10.11$; $SD = 2.66$), $t(35) = 2.07$; $p < .05$. No interaction effects were significant (all $ps > .05$).

For the *Text review access*, there was a main effect of question-answering success, $F(1, 35) = 82.15$; $p < .001$; $\eta^2 = .47$. Pairwise comparisons revealed that participants decided to access the ideal question-relevant text information more often after failure ($M = .77$; $SD = .23$) than after success ($M = .33$; $SD = .28$), $t(35) = 9.02$; $p < .001$. No other main effects or interactions were significant (all $ps > .05$).

For the *Text review reading time* (in seconds), there was a main effect of think-aloud method, $F(1, 16) = 8.65$; $p = .01$; $\eta^2 = .18$. Pairwise comparisons revealed that participants spent more time reviewing the ideal question-relevant text information when thinking-aloud ($M = 12.26$; $SD = 6.19$) than when not thinking-aloud ($M = 7.81$; $SD = 3.00$), $t(16) = 2.94$; $p = .01$. No other main effects or interactions were significant (all $ps > .05$).

Overall, as predicted, students paid more attention to feedback after failure than after success. Students spent more time processing the feedback automatically delivered and decided to access the optional ideal question-relevant text information more often after failure than after success. We also found that less-skilled comprehenders devoted more time processing the automatically-delivered feedback messages than skilled comprehenders, but there were no differences in processing the optional feedback messages. Interestingly, thinking-aloud increased the total time students spent reading the automatically-delivered feedback and the optional feedback, but did not influence the students' decision to access the text review screen.

3.3. Prevalence of cognitive and metacognitive processes during feedback processing and relation to performance

We examined the prevalence of meaning making, monitoring, and self-regulation reports during feedback processing. Before performing these analyses, we examined and confirmed that there were no differences between skilled and less-skilled comprehenders in the aggregate of meaning making, monitoring and self-regulation

reports produced ($M = 23.50$; $SD = 8.97$, and $M = 26.05$; $SD = 12.85$, for skilled and less-skilled comprehenders, respectively, $t(38) = .73$, $p > .05$). Table 4 presents the proportion for each think-aloud category reported by all 40 participants across the 10 questions answered during thinking-aloud. Results showed that 50% of students' verbal reports referred to *monitoring* processes, 37% referred to *meaning making* processes, and only 8% referred to *self-regulation* processes (.50, .37, and .08 in Table 4).

Table 4. *Descriptive Statistics of Think-Aloud Categories Reported when Processing Feedback*

Think-aloud category	Overall <i>M</i> proportion	<i>SD</i>	<i>Range</i>
Meaning Making	.37	.20	0 – .80
Restatements	.13	.15	0 – .48
Paraphrases	.16	.12	0 – .50
Elaborations	.08	.14	0 – .67
Monitoring	.50	.20	.11 – 1.00
Performance Comparison	.18	.12	0 – .50
Affective Reaction	.17	.15	0 – .58
Attribution	.09	.09	0 – .40
Assessing Compreh. & Memory	.05	.07	0 – .35
Self-regulation	.08	.09	0 – .43
Others	.05	.05	0 – .22

We examined whether the cognitive and metacognitive processes varied as a function of student's comprehension skill. Table 5 reports the means and standard deviations for the meaning making, monitoring and self-regulation categories as a function of comprehension skill. A mixed ANOVA with the *meaning making* categories (paraphrases, restatements, elaborations) as within-subjects variables and comprehension skill (less-skilled, skilled) as a between-subjects variable was conducted. There was a main effect for the proportion of meaning making verbalizations, $F(1, 38) = 6.30$; $p = .016$; $\eta^2 = .07$. Pairwise comparisons revealed that participants reported more paraphrases ($M = .16$; $SD = .12$) than elaborations ($M = .08$; $SD = .14$), $t(38) = 2.47$; $p = .016$. No other comparisons were significant (all $ps > .05$).

There was no main effect of comprehension skill, $F(1, 38) = .4$; $p > .05$, or interaction between the meaning making categories and comprehension skill level, $F(1, 38) = .62$; $p > .05$. Thus, we did not confirm our initial hypothesis that skilled comprehenders would deploy more meaning making reports compared to the less-skilled comprehenders. Results for meaning making processes were confirmed with a non-parametric Friedman test, $\chi^2(2) = 12.50$; $p = .002$. Non-parametric Wilcoxon signed rank tests showed that participants reported more paraphrases than elaborations, $Z = -2.99$; $p = .003$, and more restatements than elaborations, $Z = -2.02$; $p = .044$.

A mixed ANOVA with *monitoring* categories (performance comparison, affective reaction, attribution, assessing comprehension and memory) as within-subjects variables, and comprehension skill (less-skilled, skilled) as a between-subjects variable was conducted. There was a main effect for the proportion of monitoring verbalizations, $F(1, 38) = 44.08$; $p < .001$; $\eta^2 = .23$. Pairwise comparisons revealed that participants reported more performance comparisons ($M = .18$; $SD = .12$) than attributions ($M = .09$; $SD = .09$), $t(38) = 3.95$; $p < .001$, or assessments of comprehension and memory ($M = .05$; $SD = .07$), $t(38) = 5.57$; $p < .001$. Moreover, participants reported more affective reactions ($M = .17$; $SD = .15$) than attributions, $t(38) = 2.88$; $p = .006$, or assessments of comprehension and memory, $t(38) = 4.14$; $p < .001$. Further, participants reported more attributions than assessments of comprehension and memory, $t(38) = 2.16$; $p = .039$. There was no difference between performance comparisons and affective reactions, $t(38) = .35$; $p > .05$. No main effect of comprehension skill, $F(1, 38) = 1.08$; $p > .05$, or interaction with monitoring categories, $F(1, 38) = .1$; $p > .05$ were found. We did not confirm our hypothesis that less-skilled comprehenders would deploy more monitoring reports compared to skilled comprehenders. Results for monitoring processes were confirmed with a non-parametric Friedman test, $\chi^2(3) = 34.39$; $p < .001$. A number of non-parametric Wilcoxon signed rank tests showed that participants reported more performance comparisons than attributions ($Z = -3.40$; $p = .001$) or assessments of comprehension and memory ($Z = -4.38$; $p < .001$). Moreover, participants reported more affective reactions than attributions ($Z = -2.74$; $p = .006$) or assessments of comprehension and memory ($Z = -3.68$; $p < .001$). Further, participants reported more attributions than assessments of comprehension and memory ($Z = -2.27$; $p = .023$). However, participants reported the same amount of performance comparisons and affective reactions ($Z = -.64$; $p > .05$).

For the *self-regulation* reports, a one-way ANOVA with the students' comprehension skill (less-skilled, skilled) as a between-subjects variable was conducted. There were no significant differences with regard to the amount of self-regulation reports between less-skilled and skilled comprehenders, $F(1, 38) = .04$; $p > .05$.

Table 5. *Descriptive Statistics of the Proportions of the Meaning Making, Monitoring, and Self-regulation Categories, as a Function of Comprehension Skill (Less-skilled, Skilled)*

Think-aloud category	Comprehension Skill	
	Less-skilled	Skilled
Meaning Making		
Paraphrases	.16 (.09)	.16 (.15)
Restatements	.14 (.15)	.13 (.15)
Elaborations	.10 (.18)	.06 (.10)
Monitoring		
Performance Comparison	.18 (.12)	.18 (.12)
Affective Reaction	.16 (.15)	.18 (.15)
Attribution	.08 (.06)	.11 (.11)
Assessing Compreh. & Memory	.05 (.06)	.06 (.09)
Self-regulation	.08 (.07)	.08 (.11)

Note. *SD* in parentheses.

A correlation analysis was conducted to examine the relation between feedback processing and students' overall performance (see Table 6). The analysis showed that student's performance was negatively related to both feedback reading time ($r = -.320$; $p < .05$) and text review access ($r = -.421$; $p < .01$). These negative relations are consistent with the findings reported above showing that students spent more time processing feedback after failure than after success. There were no significant relations between performance and cognitive or metacognitive processes. Further, there was a significant positive relation between the time devoted to reading the automatically-delivered feedback messages and the time devoted to reading the ideal text selection ($r = .411$; $p < .01$). Finally, meaning making processes were negatively related to

monitoring processes ($r = -.889$; $p < .01$), perhaps suggesting that the more coherence is established, the lower the need for monitoring.

Table 6. *Pearson Correlations of Measures*

Measure	1	2	3	4	5	6	7	8
1. Q-A Performance	-	-.320*	-.421**	.020	-.063	.165	-.076	-.261
2. Feedback RT		-	.140	.411**	.020	-.087	.159	-.023
3. Text Rev. Access			-	.035	-.011	.025	-.012	-.030
4. Text Rev. RT				-	-.102	-.041	.308	.006
5. Meaning Making					-	-.889**	-.213	-.045
6. Monitoring						-	-.186	-.111
7. Self-regulation							-	-.223
8. Others								-

* $p \leq .05$. ** $p \leq .01$.

Note. Q-A Performance = number of correct responses; Feedback RT = time spent processing the automatically-delivered feedback; Text Rev. Access = proportion of decisions to access the optionally-delivered feedback on the ideal text selection; Text Rev. RT = time spent processing the ideal text selection.

4. DISCUSSION

The goal of this study was to examine how students process formative feedback that includes both corrective and elaborative information during question-answering tasks, and to explore potential individual differences between skilled and less-skilled comprehenders in feedback processing. Answering comprehension questions from written documents is one of the most frequent activities students perform in school settings. In this context, formative feedback can provide students not only with information about the correctness of a response, but also with hints, prompts and explanations about their strategies (e.g., Fox et al., 2014; Llorens et al., 2016; Timmers & Veldkamp, 2011). Although both theoretical and empirical works acknowledge that learners have to actively process the feedback messages for the benefits of feedback to be realized, little is known about what cognitive and metacognitive operations students deploy during feedback processing. Following the LP-FILE model of feedback

processing (Timms et al., 2016) and a theoretical model of question-answering (Rouet et al., 2017; Vidal-Abarca et al., 2010), we formulated and evaluated three hypotheses.

The first hypothesis predicted that students would concentrate on processing feedback that provides corrective information about the student's answer, namely knowledge of response (KR) and knowledge of correct response (KCR) feedback. The results confirmed this prediction and showed that students were mainly interested in knowing the answer correctness rather than getting information about the accuracy of their current text selection. As noted in the introduction, EF is usually delivered along with corrective feedback (Van der Kleij et al., 2015). In these situations, it is essential to examine the extent to which corrective feedback influences EF processing. Previous research showed that corrective feedback functions as a source that influences better monitoring and usage of EF when needed (e.g., Fox et al., 2014). In other words, receiving KR or KCR feedback serves to verify whether students' mental representation is correct or not, thus facilitating the need to access optional EF after providing incorrect responses. These findings suggest that students' goals during question-answering are mainly related to producing the right answer. This interpretation makes good sense with both theoretical models and empirical studies on question-answering (Farr et al., 1990; Rouet et al., 2017; Rupp et al., 2006; Vidal-Abarca et al., 2010). It is also possible that students do not pay attention to the elaborative information because they do not find it useful to perform the task; however, no data were collected to allow us to explore this possibility. Furthermore, there were no differences between skilled and less-skilled comprehenders. Thus, all students tend to focus on the correctness of their response, at least when no specific instructions for deep learning are provided. An interesting question raised by these findings is the extent to which students' goals (e.g., responding to questions to comprehend the text versus to learn the content) influence engagement in feedback processing. This would be important to address in future research.

The second hypothesis predicted that students would pay more attention to feedback after failure than after success. This prediction was in line with the LP-FILE model (Timms et al., 2016) and those in the extant literature. Indeed, students devoted more time to processing feedback after failure than after success. Interestingly, this finding was independent from the students' level of comprehension. These results suggest that knowing that an answer is incorrect likely challenges students and, in turn, leads to devoting more resources to understanding failure, whereas knowing that an

answer is correct confirms the students' previous response and reinforces their understanding (Maier et al., 2016).

The third hypothesis predicted that students would deploy cognitive and metacognitive processes closely related to question-answering performance. Regarding cognitive processes, the findings showed that students paraphrased and elaborated the feedback messages. Regarding metacognitive processes, students monitored their response (e.g., comparing their current response to the correct answer, assessing their comprehension and memory, and reporting affective reactions and attributions), and self-regulated their behaviors (e.g., verbalizations and actions aimed at remediating their mistakes).

The most prevalent monitoring process was performance comparisons, indicating that students used the feedback to compare their performance against the desired level of performance. It is important to note that these comparisons were usually reported in a very shallow way (e.g., "*Ok, it was not this one, but this one*"; "*I should have selected this information too*"), which may also reflect that students focus on the correctness of response, rather than understanding the reasons for failure. Students also reacted affectively and made attributions, which is an important component of their metacognitive processing. These findings are consistent with the control-value theory of achievement emotions (Pekrun, Frenzel, Goetz, & Perry, 2007), which predicts that feedback after learning achievement causes emotions (e.g., frustration after failure or joy after success) as well as attributions (e.g., lack of effort), which are mediated by achievement goals and internal variables (e.g., temperament).

Contrary to our predictions, no differences between skilled and less-skilled comprehenders were observed for the cognitive and metacognitive processes analyzed. This could be due, in part, to lack of sensitivity in the coding scheme of these processes. In this study, we coded both positive and negative value reports under the same label because of the low rates encountered in each coding category. For instance, '*Great, that was correct*' and '*Today is not my day*' were coded as affective reactions, although their connotations are opposite. Further, the sample was split into skilled and less-skilled comprehenders using scores from a task that closely aligns to the experimental task: answer comprehension questions from an available text. Thus, this measure evaluated reading skill as it relates to question-answering tasks (i.e., creating a task model and searching the text for question-relevant information), not general reading

comprehension performance. Thus, more work is needed to better understand whether and how reading skill influences feedback processing.

Self-regulation processes were much less frequent than monitoring processes. This finding may also be a function of the actual demands of the question-answering task. In this study, students had no option for a second attempt after automatic feedback was delivered. Having a second attempt may have increased self-regulation activities employed to improve the response. Thus, it sounds logical that students concentrated primarily on monitoring activities aimed at evaluating their current response. The absence of a second attempt to provide an answer may also explain the absence of significant relations among performance and cognitive or metacognitive processes.

The think-aloud methodology was used to investigate the students' processing of feedback. In this context, it is important to acknowledge the ongoing debate about the extent to which thinking-aloud influences the actual cognitive processes under investigation (see Fox, Ericsson, & Best, 2011, and comments by Schooler, 2011). In the present study, we employed the think-aloud methodology as a within-subjects condition for the behavioral online measures on students' processes. As expected, thinking-aloud influenced the two reading time measurements. Students spent more time overall in the think-aloud condition than in the silent reading condition. However, thinking-aloud did not influence the number of decisions to access the optional EF, nor interacted with individual differences in comprehension skill. Although thinking-aloud always has the potential to influence cognitive processing, the findings of the present study suggest that such influence was rather minimal, which is consistent with previous findings in the field of self-regulated learning (e.g., Veenman, Elshout, & Groen, 1993). Veenman et al. (1993) found correspondence between learning events recorded in computer log files and learners' think-aloud reports for metacognitive monitoring and control processes despite the fact that thinking-aloud increased time on task.

It is also important to note that students performed the question-answering task in a digital environment rather than a traditional non-digital paper-and-pencil environment. A recent meta-analysis showed that reading in digital media hinders text comprehension compared to reading in printed media, specifically for expository texts (Delgado, Vargas, Ackerman, & Salmerón, 2018). Poor comprehension and achievement monitoring may explain this effect (e.g., Ackerman & Goldsmith, 2011). For example, learners tend to overestimate their level of learning in digital media (Ackerman & Lauterman, 2012). As a result, they may not engage in processing the text

information deeply, which may be detrimental to both comprehension and learning. Detrimental effects of this shallow-level processing may be amplified due to increased demands on decision-making processes in digital environments (e.g., moving from one screen to another; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018). Even though these findings raise questions about the potential effects of the digital environment in the present study, the absence of a non-digital print comparison condition precludes any conclusions to be drawn. Furthermore, such a comparison may not be feasible at all in this context. Providing students with immediate, complex elaborated feedback in a paper-and-pencil formative assessment would require evaluating students' responses in real time and one-to-one tutoring.

The findings of this study have important implications for the theory and practice of formative feedback. Theoretical models of feedback, such as the LP-FILE model (Timms et al., 2016), could be enriched by incorporating the cognitive and metacognitive processes observed as part of the active feedback processing they posit. From a practical point of view, the findings have implications for the design of educational digital environments (van den Broek, Kendeou, & White, 2009). For example, the specific feedback administered in the present study has been found to enhance students' comprehension and improve skills to search and evaluate textual relevance (Máñez et al., 2016; Llorens et al., 2016), which suggests that digital environments may benefit from elaborative feedback that provides corrective information along with hints on how to proceed. However, the findings also suggest that students may not be willing to spend a lot of time processing feedback information. Thus, designers should consider carefully how much feedback information they provide and implement procedures that encourage students to use feedback when they disregard it.

The present study also has a number of limitations that present opportunities for future research. First, participants performed a relatively low-stakes question-answering task that may be very frequent in school settings, but is not very demanding. These low task demands may have led students to focus their attention on their answer correctness rather than the elaborative feedback related to the selection task. In this context, perhaps complex formative feedback about specific strategies may be more effective at focusing attention on feedback than the elaborative feedback we employed. Also, including a scoring system that rewards correct responses may enhance students' engagement in processing elaborative components of feedback. Second, participants' individual

differences in prior knowledge is an important factor (Kendeou & O'Brien, 2016) we have not considered. For example, empirical evidence in the area of mathematics suggests that primary school students and university students with little prior knowledge benefit from feedback interventions more than participants with higher prior knowledge (e.g., Fyfe, Rittle-Johnson, & DeCaro, 2012; Krause et al., 2009). Even though we controlled to some extent demands of prior knowledge with the inclusion of non-academic topic expository texts, more research is needed to understand how prior knowledge influences formative feedback processing. Third, the automatically-delivered feedback included multiple components (i.e., KR, KCR and EF regarding the relevant information selected), which may have increased cognitive load. The effects obtained cannot be attributed to individual feedback components, but rather to additive or interactive effects among these components (Jordan, 2012). Future investigations can isolate and examine these components to tease their specific effects.

In conclusion, the findings of the present study contribute to our understanding of how students process complex formative feedback in digital environments that include answering questions from an available text. This study reveals that students pay attention to corrective feedback information frequently, but they rarely engage in processing more elaborative feedback. Furthermore, their attention to feedback is mostly driven by initial failure. The findings also highlight the prevalence of both cognitive and metacognitive operations students deploy while processing of feedback. Finally, this research suggests that comprehension skill (at least as operationalized in the present study) has little to no effect on how students engage in computer-based feedback processing.

References

- Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, *17*(1), 18-32. doi:10.1037/a0022086
- Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, *28*(5), 1816-1828. doi:10.1016/j.chb.2012.04.023
- 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
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*(3), 206-219. doi:10.1016/j.cedpsych.2014.05.003
- Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, *100*(1), 209–222. doi:10.1037/0022-0663.100.1.209
- Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011). Selecting information to answer questions: Strategic individual differences when searching texts. *Learning and Individual Differences*, *21*(2), 201–205. doi:10.1016/j.lindif.2010.11.007
- Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2013). Self-generated explanations on the question demands are not always helpful. *The Spanish Journal of Psychology*, *16*(e26), 1-11. doi:10.1017/sjp.2013.45
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading

- comprehension. *Educational Research Review*, 25, 23-38.
doi:10.1016/j.edurev.2018.09.003
- Dunlosky, J., & Metcalfe, J. (2009). *Metacognition: A textbook for cognitive, educational, life span and applied psychology*. Thousand Oaks, CA: Sage.
- Farr, R., Pritchard, R., & Smitten, B. (1990). A description of what happens when an examinee takes a multiple-choice reading comprehension test. *Journal of Educational Measurement*, 27(3), 209-226. doi:10.1111/j.1745-3984.1990.tb00744.x
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316-344. doi:10.1037/a0021663
- Fox, J.-P., Klein Entink, R., & Timmers, C. (2014). The joint multivariate modeling of multiple mixed response sources: Relating student performances with feedback behavior. *Multivariate Behavioral Research*, 49(1), 54-66.
doi:10.1080/00273171.2013.843441
- Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104(4), 1094–1108. doi:10.1037/a0028389
- Graesser, A. C., & Murachver, T. (1985). Symbolic procedures of question answering. In A. C. Graesser & J. B. Black (Eds.), *The psychology of questions* (pp. 15-88). Hillsdale, NJ: Erlbaum.
- Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249-271). New York, NY: Routledge.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Jordan, S. (2012). Student engagement with assessment and feedback: some lessons from short-answer free-text e-assessment questions. *Computers & Education*, 58(2), 818-834. doi:10.1016/j.compedu.2011.10.007
- Kendeou, P., & O'Brien, E. J. (2016). Prior knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text and context* (pp. 151-163). New York, NY: Routledge.
- Kendeou, P., McMaster, K. L., & Christ, T. J. (2016). Reading comprehension: Core components and processes. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 62-69. doi:10.1177/2372732215624707
- Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, 19(2), 158 –170. doi:10.1016/j.learninstruc.2008.03.003
- 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., & 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
- 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

- Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*(3), 251-283. doi:10.1207/S1532690XCI2103_02
- Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education, 95*, 85-98. doi:10.1016/j.compedu.2015.12.002
- Máñez, I., Vidal-Abarca, E., & Martínez, T. (2016, July). Accuracy to select relevant text information to answer questions from a text. Paper presented at the 26th annual meeting of the Society for Text and Discourse, Kassel, Germany.
- 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. doi:10.1174/021037008785702956
- Martínez, T., Vidal-Abarca, E., Gil, L., & Gilabert, R. (2009). On-line assessment of comprehension processes. *The Spanish Journal of Psychology, 12*(1), 308-319. doi:10.1017/S1138741600001700
- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*, 99–113. doi:10.1023/B:TRUC.0000021811.66966.1d
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Erlbaum.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research*

- on educational communications and technology* (pp. 125–144). New York, NY: Erlbaum.
- Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brünken (Eds.), *Instructional design for multimedia learning* (pp. 181-195). Münster: Waxmann.
- Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Gogvadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education, 71*, 56-76. doi:10.1016/j.compedu.2013.09.011
- Ness, M. (2011). Explicit reading comprehension instruction in elementary classrooms: Teacher use of reading comprehension strategies. *Journal of Research in Childhood Education, 25*(1), 98–117. doi:10.1080/02568543.2010.531076
- Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz, & R. Pekrun, (Eds.), *Emotion in education* (pp. 13-36). San Diego, CA: Academic Press.
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.
- Pressley, M., Wharton-McDonald, R., Mistretta-Hampston, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth-grade classrooms in upstate New York. *Scientific Studies of Reading, 2*(2), 159-194. doi:10.1207/s1532799xssr0202_4
- Ramos, J. L., & Cuetos, F. (1999). *PROLEC-SE: Evaluación de los procesos lectores en alumnos de tercer ciclo de educación primaria y secundaria*. Madrid: TEA Ediciones.

- 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
- Roelle, J. & Berthold. K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, 49(2), 142-156. doi:10.1016/j.learninstruc.2017.01.008
- Rouet, J-F., Britt, M. A., & Durik, A. M., (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200-215. doi:10.1080/00461520.2017.1329015
- Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, 23(4), 441–474. doi:10.1191/0265532206lt337oa
- Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadler, M., & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91-120). Amsterdam, The Netherlands: John Benjamins.
- Schooler, J. W. (2011). Introspecting in the spirit of William James: Comment on Fox, Ericsson, and Best (2011). *Psychological Bulletin*, 137(2), 345-350. doi:10.1037/a0022390
- 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(3), 923–930. doi:10.1016/j.compedu.2010.11.007

- Timms, M., DeVelle, S., & Lay, D. (2016). Towards a model of how learners process feedback: A deeper look at learning. *Australian Journal of Education*, 60(2), 128-145. doi:10.1177/0004944116652912
- van den Broek, P., Kendeou, P., & White, M. J. (2009). Cognitive processes during reading: Implications for the use of multimedia to foster reading comprehension. In A. G. Bus, & S. B. Neuman (Eds.), *Multimedia and literacy development: Improving achievement for young learners* (pp. 57-73). New York, NY: Routledge.
- Van der Kleij, F. M., Eggen, T. J. H. M., Timmers, C. F., & Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, 58(1), 263–272. doi:10.1016/j.compedu.2011.07.020
- Van der Kleij, F. M., Feskens, R. C. W., & 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
- Veenman, M. V. J., Elshout, J. J., & Groen, M. G. M. (1993). Thinking aloud: Does it affect regulatory processes in learning? *Tijdschrift voor Onderwijsresearch*, 18(6), 322-330.
- Vidal-Abarca, E., Gilabert, R., & Abad, N. (2002a). 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., Reyes, H., Gilabert, R., Calpe, J., Soria, E., & Graesser, A. C. (2002b). ETAT: Expository Text Analysis Tool. *Behavior Research Methods, Instruments, & Computers*, 34(1), 93-107. doi:10.3758/BF03195428

- 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. C., & Ferris, R. (2011). Recording online processes in task-oriented reading with Read&Answer. *Behavior Research Methods*, *43*(1), 179–192. doi:10.3758/s13428-010-0032-1
- Wolfe, M. B., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*(4), 467-502. doi:10.1207/s1532690xci2304_2

CAPÍTULO 4

ESTUDIO 3: ¿INFLUYE LA RETROALIMENTACIÓN CORRECTIVA EN EL USO DE LA RETROALIMENTACIÓN ELABORADA EN UN ENTORNO DIGITAL?

TITLE: DOES CORRECTIVE FEEDBACK INFLUENCE THE USE OF ELABORATED FEEDBACK IN A DIGITAL ENVIRONMENT?

Manuscript submitted for publication

RESUMEN

Los entornos digitales permiten proporcionar retroalimentación correctiva de verificación de respuestas (KR: *Correcto/Incorrecto*) o de respuesta correcta (KCR: *La respuesta correcta es X*), junto con retroalimentación elaborada (EF: Explicaciones). Es frecuente asumir que los estudiantes procesan la EF, aunque investigaciones recientes demuestran que eso no siempre es así, debido probablemente a la retroalimentación correctiva proporcionada junto con la EF. El objetivo de este trabajo es analizar la decisión voluntaria de acceder a EF en función de la retroalimentación KR y KCR recibida inmediatamente tras responder preguntas de un texto de ciencias. Estudiantes de Secundaria recibieron retroalimentación correctiva según la condición asignada: KR, KCR, o Control (i.e., no correctiva). Posteriormente, todos pudieron acceder a EF que incluía una explicación sobre los conocimientos evaluados. Los resultados sugieren que recibir retroalimentación KR y KCR disminuye el uso de EF en comparación con la retroalimentación Control, aunque no hay diferencias en el aprendizaje final.

Palabras clave: retroalimentación correctiva; retroalimentación elaborada; aprendizaje; entornos digitales

ABSTRACT

Digital environments are able to provide corrective feedback such as Knowledge of Response (KR: *Correct/Incorrect*) or Knowledge of Correct Response (KCR: *The correct answer is X*), along with Elaborated Feedback (EF: Explanations). It is common to assume that students process EF, although recent research shows that this is not always the case, probably because of the corrective feedback provided along with the EF. The main goal of this study is to analyze the voluntary decision to access EF depending on the KR and KCR feedback received immediately after answering questions from a science text. Secondary students received corrective feedback according to the condition assigned: KR, KCR, or Control (i.e., non-corrective feedback). Afterwards, all the students could access EF that included an explanation about the knowledge assessed. Results suggest that receiving KR and KCR feedback decreases the use of EF compared to the Control feedback, although there are no differences in the final learning.

Keywords: corrective feedback; elaborated feedback; learning; digital environments

EXTENDED SUMMARY

Answering questions from a text and receiving formative feedback is a common academic task (Ness, 2011) that can be delivered in digital environments (Shute & Rahimi, 2017). Formative feedback refers to any information provided to the students about their performance to improve their learning (Shute, 2008). Formative feedback embraces a range of feedback types. A common distinction is usually made between corrective and elaborative feedback. Corrective feedback provides students with Knowledge of Response or KR (*Correct/Incorrect*) or Knowledge of Correct Response or KCR (*The correct answer is X*). Elaborated Feedback or EF usually contains additional information beyond student's performance such as explanations or hints. A recent review study suggests that EF is the most effective type of feedback to enhance learning in digital environments, KCR feedback has a lower effect, and KR does not influence learning outcomes (van der Kleij, Feskens, & Eggen, 2015).

In order to be effective, students have to engage actively in processing its content (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). However, little research has explored how students process feedback in digital environments. Recent findings suggest that students tend to focus their attention on knowing the accuracy of their responses and the correct answer even when EF is received (Lefevre & Cox, 2016; Máñez, Vidal-Abarca, Kendeou, & Martínez, 2019). Although EF is usually provided along with corrective feedback, little is known about the effect of receiving KR and KCR feedback on the student's decision to use EF. Thus, our study examines the extent to which corrective KR and KCR feedback may influence the student's decision to voluntarily access EF when answering questions from a science text in a digital environment. Likewise, this study also examines whether these types of feedback influence the student's performance when learning conceptual knowledge in the area of physics.

Method

Sixty-seven 9th-grade students ($M_{\text{age}} = 15$; $SD = .82$) participated in the study (52.2% were male and 47.8% were female). All participants were Spanish native speakers from two schools of Valencia (Spain). All the students had studied Natural Sciences and Physics according to the Spanish curriculum. After measuring student's prior knowledge, each participant was assigned to one of the three experimental groups that varied in the type of corrective feedback delivered: KR, KCR, or Control (i.e., non-corrective feedback). For the learning task, students read a science text about the

atmospheric pressure and wind phenomenon and answered a set of multiple-choice questions with the text available in Read&Learn, a web-based application that traces students' behaviors. After answering each question, Read&Learn delivered corrective feedback automatically depending on the experimental condition assigned. Whereas students in the KR group received information on their response correctness (*Correct/Incorrect*), students in the KCR group were informed about the correct response (*The correct answer is X*), and students in the Control group received non-corrective feedback (e.g., *You have answered question number X*). After receiving these feedback messages, all the students had the option to voluntarily access EF made of an explanation to infer the correct response without stating it explicitly. For instance, for the question '*Is the air density the same in all points of the earth?*' (Correct response: *No, because it varies depending on the temperature and the altitude*), the EF message is '*Think that the air density is determined by how close or separated the particles are, and that depends on the external conditions*'. Finally, participants answered a final task with new short-answer questions 24 h later. Read&Learn recorded the student's decisions to access EF, which were conditioned to question-answering success. We computed two performance scores: learning task performance and final task performance as the number of correct responses.

Results

To address the first aim, to analyze the effect of corrective feedback (KR, KCR, and Control) on the use of EF, we conducted Kruskal-Wallis tests for both the percentage of students who accessed the EF and the percentage of EF accesses. Significant differences were found for both measures as a function of the type of corrective feedback received. The Jonckheere tests revealed a significant trend in the data, indicating that as the corrective feedback includes more specific information about the response (KR and KCR feedback), the use of EF decreases significantly. That is, less students decided to access EF and lower rates of accesses appeared, especially when corrective feedback included the correct response (i.e., KCR). To better understand the effect of the type of corrective feedback on the use of EF, we took into account the question-answering success (correct vs. incorrect responses). The Wilcoxon tests revealed that students in the Control group accessed EF in a similar fashion after providing either correct or incorrect responses. However, students in the KR and KCR feedback groups accessed EF more frequently after providing incorrect responses (see Figure 4).

To address our second aim, to examine the effect of feedback type on the students' performance, we conducted two ANCOVAs for both learning task performance and final performance. The student's level of prior knowledge was introduced as covariate. No significant differences were found neither for the learning task performance nor the final task performance as a function of feedback type. The students' prior knowledge significantly explained both performance measures.

Discussion

It is common to assume that students process EF when received. Recent research suggests that this is not always the case, probably because of the corrective feedback provided along with the EF. Regarding our first goal, results showed a trend in the EF use based on the corrective feedback provided. Providing KR and KCR feedback seems to discourage students from accessing EF to some extent, probably because many students focus on knowing the answer correctness (Máñez et al., 2019). Students who do not know if their answers are correct tend to access EF quite often, probably looking for clues to confirm their response model or to uncover the correct answer. Receiving KR feedback triggers lower rates of EF accesses, since students tend to use EF after failure. Students who received KCR feedback seldom access EF, probably because they consider unnecessary to process additional information when they already know the correct response. Moreover, results indicate that KR and KCR feedback's main function is to verify the student's response model and hence increase the use of EF after failure (Fox, Klein Entink, & Timmers, 2014), since students struggle to monitor the accuracy of their responses. Regarding our second goal, results revealed no differences among the feedback types, which is consistent with recent findings in the areas of text comprehension and science learning (Golke, Dörfler, & Artelt, 2015; Maier, Wolf, & Randler, 2016). This result may be explained based on the EF use each experimental group made. The greater use of EF in the Control and KR feedback groups may have compensated for the positive effect of knowing the correct answer in the KCR group. Overall, findings suggest that students access EF in an effort to know the correct answer since they use EF as a function of the type of corrective feedback in combination with the accuracy of their responses. This study contributes to our understanding of how Secondary-school students engage in processing formative feedback when learning conceptual knowledge and provides valuable information for the design of digital learning environments.

1. INTRODUCCIÓN

La retroalimentación formativa es un procedimiento instruccional consistente en ofrecer información sobre el rendimiento de los estudiantes para mejorar su aprendizaje (Narciss et al., 2014; Shute, 2008). El diseño de entornos digitales de aprendizaje permite proporcionar retroalimentación formativa en tiempo real. Sin embargo, hay muy poca investigación sobre el uso que los estudiantes hacen de esta información. Un problema añadido es la variabilidad de tipos de retroalimentación existente. Siguiendo a Narciss (2013) y Shute (2008) podemos diferenciar entre retroalimentación de verificación (Knowledge of Response o KR), de respuesta correcta (Knowledge of Correct Response o KCR), y retroalimentación elaborada (Elaborated Feedback o EF). La retroalimentación KR incluye exclusivamente información de verificación de respuestas (*Correcto* o *Incorrecto*), mientras que la retroalimentación KCR incluye además la respuesta correcta (*La alternativa X es la solución correcta*), y la retroalimentación EF incluye información adicional como, por ejemplo, explicaciones sobre la respuesta. Aunque EF es el tipo de retroalimentación más efectivo en entornos digitales (van der Kleij, Feskens y Eggen, 2015), sus efectos no son automáticos, sino que los aprendices deben decidir procesar activamente su información (Bangert-Drowns, Kulik, Kulik y Morgan, 1991). EF suele proporcionarse junto a retroalimentación KR o KCR, aunque no siempre es así (p.e., se pueden dar explicaciones o pistas sobre por qué una respuesta es correcta o incorrecta, sin mencionar explícitamente la respuesta correcta). Por tanto, cabe preguntarse en qué medida la retroalimentación KR y KCR afecta a la decisión de acceder voluntariamente a EF. Éste es el principal objetivo del presente estudio. De forma complementaria, este estudio también examina el efecto de estos tipos de retroalimentación en el aprendizaje de conocimientos conceptuales en el área de física.

Habitualmente, los profesores asignan a sus alumnos tareas en las que han de responder preguntas sobre los contenidos de uno o varios textos (Ness, 2011). La razón es que estas tareas inducen procesos de comprensión que facilitan el aprendizaje (Anmarkrud, McCrudden, Bråten y Strømsø, 2013; Cerdán, Gilabert y Vidal-Abarca, 2011; Roelle y Berthold, 2017). Además de estos procesos cognitivos de comprensión (p.e., hacer inferencias), responder preguntas de un texto también implica activar procesos metacognitivos propios de la resolución de problemas (p.e., búsqueda de información en los materiales o evaluación de la relevancia de la información

encontrada) (Rouet, Britt y Durik, 2017). Un estudiante competente posee habilidades cognitivas y metacognitivas para poner en marcha esas operaciones cognitivas de comprensión, búsqueda de información y monitorización de información relevante que le conducirá a completar con éxito la tarea de responder preguntas y a aprender los conceptos o ideas requeridos (Vidal-Abarca, Mañá y Gil, 2010). Sin embargo, muchos estudiantes de Educación Secundaria suelen experimentar dificultades en este tipo de tareas (OECD, 2006 y 2010). Esa puede ser una de las razones de que los profesores proporcionen retroalimentación sobre las respuestas de sus alumnos (Ness, 2011). Seguramente asumen que la retroalimentación puede ayudar a superar las dificultades de comprensión y aprendizaje de conocimientos de ciencias de sus estudiantes. Este procedimiento instruccional, i.e., leer textos y responder a preguntas, se adapta perfectamente a los entornos digitales de aprendizaje, en los que una retroalimentación inmediata puede jugar un papel esencial (Shute y Rahimi, 2017; van der Kleij et al., 2015).

Resultados de un reciente meta-análisis sobre la efectividad de la retroalimentación en entornos digitales sugieren que EF es el tipo de retroalimentación más eficaz, mientras KCR tiene un efecto positivo menor, y KR apenas tiene efectos en el aprendizaje (van der Kleij et al., 2015). Sin embargo, la investigación previa no es concluyente por lo que respecta a la efectividad de la EF en tareas de comprensión que requieren contestar preguntas de un texto (Golke, Dörfler y Artelt, 2015; Llorens, Vidal-Abarca y Cerdán, 2016; Llorens, Vidal-Abarca, Cerdán y Ávila, 2015). Por ejemplo, Golke et al. (2015) examinaron la efectividad de tres tipos diferentes de EF diseñados para mejorar la comprensión de textos: KR + pistas para facilitar la construcción de inferencias, KR + explicaciones del error, o KR + apoyo para monitorizar la comprensión del texto. La efectividad de estos tipos de EF fue comparada con KR y con una condición control sin retroalimentación. Estudiantes de sexto grado de Primaria leyeron varios textos expositivos y narrativos para responder preguntas de comprensión tipo test en un entorno digital. Todos los estudiantes, excepto los del grupo sin retroalimentación, recibieron los mensajes asignados después de contestar incorrectamente, tras lo cual tuvieron un segundo intento sin ningún tipo de retroalimentación. Los estudiantes completaron una evaluación inmediatamente después de finalizar la prueba y otra varias semanas después. Ambas pruebas incluyeron diferentes textos y preguntas. Los resultados mostraron que no existían diferencias entre los tipos de retroalimentación EF en comparación con la retroalimentación KR o el

grupo sin retroalimentación. Además, la retroalimentación KR no influyó tampoco en el rendimiento de los estudiantes. Resultados paralelos fueron encontrados en su segundo estudio. Los autores concluyeron que la ausencia de efectos de la EF en la comprensión de los textos se debió probablemente al bajo compromiso de los estudiantes en procesar activamente la retroalimentación proporcionada. Sin embargo, estos autores no examinaron hasta qué punto los estudiantes decidieron utilizar la retroalimentación elaborada (EF), aspecto central de nuestro trabajo.

En el área de ciencias, Maier, Wolf y Randler (2016) realizaron un estudio para analizar la efectividad de varios tipos de retroalimentación: EF (KR + explicación tras responder incorrectamente), KR, o control sin retroalimentación. Estudiantes de secundaria leyeron un texto de biología, contestaron una serie de preguntas y recibieron retroalimentación según el grupo asignado. Además, pidieron a los alumnos del grupo EF que autoevaluaran el uso que habían realizado de esa información. Los resultados mostraron que, en general, no hubo diferencias entre los tres tipos de retroalimentación en el rendimiento de los estudiantes. Posteriormente, Maier et al. (2016) dividieron a los estudiantes del grupo EF en dos subgrupos: aquellos que indicaron haber utilizado la EF y aquellos que no. Los nuevos resultados mostraron que aquellos estudiantes que decían haber utilizado la EF obtuvieron mejor rendimiento que los estudiantes del grupo control y aquellos que no habían utilizado la EF. No obstante, este estudio cuenta con una limitación obvia al no utilizar una medida objetiva sobre el uso de EF. En el presente trabajo examinamos la decisión de los estudiantes de acceder voluntariamente a la EF en función de la retroalimentación correctiva recibida, así como su influencia en el aprendizaje de conocimientos conceptuales de física sobre la presión atmosférica y el viento. En nuestro estudio, los mensajes EF para cada pregunta consistían en una explicación para facilitar la comprensión de los conocimientos evaluados. Por ejemplo, para la pregunta *‘Si Torricelli hubiera cogido un tubo de 2 metros de alto lleno de mercurio, en vez de un tubo de un metro, y hubiera repetido su experimento, ¿qué habría pasado con la columna de mercurio?’*, la respuesta correcta sería *‘Habría quedado a la misma altura, 760 mm’*. En este caso, el mensaje EF era: *‘Piensa que la altura a la que queda el mercurio está determinada por el punto en el que se iguala el peso del mercurio que contiene el tubo y el peso de la columna de aire sobre la cubeta’*. Así, este mensaje contiene una explicación para facilitar la comprensión del experimento de Torricelli para demostrar la presión atmosférica, aunque sin proporcionar la respuesta correcta o KCR.

Aunque la efectividad de la retroalimentación en entornos digitales ha sido ampliamente investigada (p.e., Hattie y Gan, 2011; Shute, 2008; van der Kleij, Eggen, Timmers, y Veldkamp, 2012), existe poca investigación acerca de cuándo y cómo los estudiantes deciden procesar la retroalimentación (Butler y Winne, 1995). Por tanto, cabe preguntarse si proporcionar retroalimentación correctiva KR o KCR influye en la decisión de acceder a EF, una cuestión esencial para el diseño de entornos digitales. Parece lógico pensar que la retroalimentación correctiva permite al estudiante confirmar su modelo de respuesta cuando responde correctamente, mientras que le induce a iniciar un proceso de corrección de errores cuando responde incorrectamente, modificando así el modelo de respuesta inicialmente construido. Los estudios realizados por Timmers y Veldkamp (2011), y Fox, Klein Entink y Timmers (2014) analizaron el uso que hacían un grupo de estudiantes universitarios de la EF después de completar un test sobre competencia en la búsqueda de información en un entorno digital. Después de responder una serie de preguntas, el sistema proporcionó retroalimentación KR para cada respuesta. Una vez se mostró la retroalimentación KR, los estudiantes pudieron acceder opcionalmente a un mensaje EF para cada pregunta. Es importante señalar que estos mensajes EF contenían retroalimentación KCR junto con una explicación. Los resultados mostraron que los estudiantes accedieron a esos mensajes EF en las preguntas respondidas incorrectamente. Una limitación importante de estos estudios es la presencia de retroalimentación KCR junto con la explicación, por lo que no podemos saber si los estudiantes accedieron a esos mensajes para conocer la respuesta correcta o también para leer la explicación.

Por otra parte, existe investigación que ha examinado el tipo de retroalimentación habitualmente proporcionado en entornos digitales de aprendizaje, así como las preferencias de los aprendices por un tipo de retroalimentación u otro. En un estudio reciente, Tärning (2018) encontró que las aplicaciones utilizadas en Educación Primaria con escolares suecos proporcionaban principalmente retroalimentación KR, y en muy pocas ocasiones proporcionaban retroalimentación KCR y EF. Como recomendación, la autora sugiere que los entornos digitales deben incluir EF con explicaciones cuando su objetivo sea facilitar el aprendizaje. Sugerencias similares fueron propuestas por Lefevre y Cox (2016), quienes examinaron el tipo de retroalimentación que preferían los estudiantes de posgrado después de responder preguntas tipo test en un entorno digital. Estos autores sugieren que los entornos digitales deben a) proporcionar retroalimentación, b) ofrecer EF para las respuestas

correctas, c) proporcionar retroalimentación clara y focalizada en la tarea, d) ofrecer EF de tipo explicativo en las preguntas que evalúan la aplicación del material estudiado, y e) permitir a los estudiantes revisar las preguntas. Los resultados de este estudio mostraron que los estudiantes prestan especial atención a la retroalimentación KR inicialmente y a la retroalimentación KCR cuando responden incorrectamente, por lo que la precisión de las respuestas es un factor esencial para comprender el uso de la EF (Lefevre y Cox, 2016). En un estudio con la metodología pensar-en-voz-alta, Máñez, Vidal-Abarca, Kendeou y Martínez (2019) analizaron el procesamiento de una EF que contenía KR, KCR y pistas para mejorar la comprensión de los contenidos evaluados. Los resultados indicaron que los estudiantes focalizan su atención en conocer la precisión de su respuesta y la respuesta correcta. No obstante lo anterior, existe un vacío respecto al papel que juega la retroalimentación correctiva KR y KCR en el uso de la EF en entornos digitales, cuestión abordada en el presente estudio.

Así, el objetivo de este estudio es, en primer lugar, analizar la influencia de la retroalimentación correctiva KR y KCR en el uso de la EF en un entorno digital. Para ello diseñamos un experimento en el que manipulamos el tipo de retroalimentación correctiva que proporcionaba el sistema automáticamente tras responder cada pregunta: retroalimentación de verificación (KR), de respuesta correcta (KCR), o no correctiva (Control). Tras recibir ese mensaje, todos los estudiantes de las tres condiciones podían acceder voluntariamente a un mensaje de retroalimentación elaborada (EF) para cada pregunta. De esta manera examinamos la influencia de la retroalimentación KR y KCR en comparación con una retroalimentación Control en la decisión de acceder a la EF. Esperamos que la retroalimentación correctiva influya en la tendencia de los estudiantes a utilizar la EF. Por tanto, el grupo Control será el que empleará la EF en mayor medida, debido a que ésta es la única fuente de información con la que comprobar la comprensión de los conocimientos evaluados. El grupo KR, sin embargo, utilizará la EF un poco menos, debido probablemente a que emplearán la EF tras responder incorrectamente. Por último, el grupo KCR será el que utilice la EF en menor medida, ya que conocer la respuesta correcta puede interferir con el uso de información adicional contenida en la EF. Además, esperamos que los estudiantes utilicen la EF tras responder incorrectamente cuando reciban retroalimentación KR y KCR, mientras que tendrán dificultades para monitorizar su uso cuando desconozcan la precisión de sus respuestas. Asimismo, este estudio examina, en segundo lugar, el efecto del tipo de retroalimentación en el rendimiento de los estudiantes en la prueba de aprendizaje y en

una prueba final administrada 24 horas después. Dos posibles predicciones pueden ser planteadas con este diseño. En primer lugar, es posible que el grupo que recibe retroalimentación KCR obtenga un rendimiento superior en ambas tareas, debido probablemente a la rápida codificación de las respuestas correctas proporcionadas. En segundo lugar, es posible que el mayor uso de la EF por parte de los estudiantes de los grupos Control y KR cancele el efecto positivo de conocer la respuesta correcta en el grupo KCR.

2. MÉTODO

2.1. Participantes

Sesenta y siete estudiantes de 3º de ESO ($M_{\text{edad}} = 15$; $DT = .82$) de dos centros educativos públicos de la Comunidad Valenciana participaron en este estudio. El 52.2% de los estudiantes eran varones y el 47.8% mujeres. Todos los alumnos habían estudiado Ciencias Naturales y Física según el plan de estudios español. La participación fue aprobada por el consejo del centro, y la recogida y tratamiento de datos fue realizada de forma que se preservó el anonimato de los participantes. Cada estudiante fue asignado a uno de los tres tipos de retroalimentación correctiva (KR, KCR, o Control), asegurándonos de que el nivel de conocimiento previo fuera similar en los tres grupos, $F(2, 64) = 0.23$; $p > .05$.

2.2. Materiales

Los materiales incluyeron un cuestionario de conocimientos previos, materiales para la tarea de aprendizaje (un texto de ciencias, preguntas de comprensión y mensajes de retroalimentación para cada pregunta), y una prueba final con nuevas preguntas. Estos materiales fueron previamente validados en dos grupos de 3º de ESO de un centro educativo público. Una vez introducidas las modificaciones oportunas, estos materiales fueron utilizados en el presente estudio.

En primer lugar evaluamos el conocimiento previo en ciencias con un cuestionario compuesto por 30 afirmaciones. Para cada afirmación, los alumnos debían indicar ‘Verdadero’, ‘Falso’ o ‘No sé’. El cuestionario incluyó conocimiento general de ciencias (*La temperatura es una propiedad de la materia*) y conocimientos específicos de la temática (*Cuando la temperatura de un gas disminuye, su presión disminuye*).

Cada respuesta correcta obtuvo 1 punto y cada incorrecta un cero; las respuestas ‘No sé’ se contaron por separado.

Toda la tarea de aprendizaje se efectuó con una aplicación informática cuyo funcionamiento se explica en la sección siguiente. Se empleó un texto sobre la presión atmosférica y el viento adaptado de un libro de texto de ciencias. El texto tenía una longitud de 1126 palabras, incluidos los títulos de los (sub)apartados y las notas de sus figuras. Para promover la comprensión y el aprendizaje de los contenidos elaboramos 12 preguntas que podían ser contestadas consultando la información del texto. Por ejemplo, *¿La densidad del aire es igual en todos los puntos de la tierra?* Para contestar esta pregunta, los alumnos debían integrar información de varias frases del texto. En este caso, los estudiantes debían comprender que la atmósfera es un espacio extremadamente grande formado por partículas cuya distribución no es uniforme, que la densidad de las capas del aire varía según la altura, y que la presión atmosférica también varía según la temperatura. Por tanto, la respuesta correcta era *No, porque varía en función de la temperatura y la altura*. Cada pregunta tenía asociado un mensaje EF que consistía en una explicación para comprender la respuesta correcta, pero sin decirlo explícitamente. Es decir, se proporcionaban únicamente pistas sobre la respuesta correcta. Cada mensaje comenzaba con la expresión *Piensa que...* seguida de la explicación pertinente sobre los fenómenos atmosféricos evaluados en la pregunta. Por ejemplo, el mensaje para la pregunta anterior fue: *Piensa que la densidad del aire se determina por lo juntas o separadas que se encuentran las partículas que lo componen, y eso depende de las condiciones externas*. Es importante señalar que los alumnos tomaban la decisión de acceder voluntariamente a estos mensajes EF después de contestar cada pregunta, siendo cada decisión registrada automáticamente por la aplicación informática. Cada respuesta fue calificada con 1 o 0 dependiendo de si era correcta o incorrecta.

La prueba final estuvo formada por 10 preguntas nuevas de respuesta breve. Estas preguntas evaluaban los conocimientos aprendidos en la tarea de aprendizaje. Por ejemplo, *¿Qué factores hacen variar la presión atmosférica?* Cada respuesta se calificó con 1 (correcta), 0,5 (parcialmente correcta) o 0 (incorrecta) en función del grado de exactitud. Para la pregunta anterior, si el estudiante responde *La altura y la temperatura* recibiría un punto, mientras que si únicamente menciona un factor recibiría 0,5 puntos.

2.3. Herramienta

La aplicación informática Read&Learn fue utilizada para la tarea de aprendizaje. Read&Learn es un desarrollo de una aplicación anterior llamada Read&Answer (Vidal-Abarca et al., 2011). Read&Learn presenta información, registra y computa las acciones y los tiempos de procesamiento de los estudiantes mientras realizan la tarea (p.e., contestar preguntas de un texto). Además, Read&Learn puede proporcionar diferentes tipos de retroalimentación adaptados a la ejecución de los estudiantes. En este estudio, Read&Learn presentaba el texto en una pantalla (Figura 1a) y la pregunta y sus alternativas de respuesta se presentaron en otra pantalla (Figura 1b). Para contestar, los estudiantes marcaban la alternativa de respuesta que consideraban correcta. Un interfaz sencillo permitía ir de una pantalla a la otra clicando el botón correspondiente en cada pantalla.

a)

Unidad - Mi aplicación A x

No es seguro | lectura.uv.es/ReadAndLearn/PL3_Experimentos/PL3_Texto?GrupoID=13588;ModuloID=2304&textoActual=0&inicioTexto=True

Preguntas

LA PRESIÓN ATMOSFÉRICA Y EL VIENTO

La Tierra está rodeada por una capa de gases que la separa del espacio vacío que constituye, en su mayor parte, el Universo. Esta capa recibe el nombre de atmósfera y está formada por una mezcla de gases que llamamos aire.

En esta unidad aprenderás algunas características básicas de la atmósfera, entre ellas, qué es presión atmosférica y uno de los fenómenos más importantes relacionados con esta: el viento.

La presión atmosférica

La atmósfera está formada por un conjunto de gases que, como toda materia, está compuesto por partículas que tienen masa y, por lo tanto, peso. Se estima que el aire que compone la atmósfera pesa, aproximadamente, 5.500 billones de toneladas.

La presión atmosférica es la fuerza que se ejerce, en un punto concreto, por el peso de la columna de aire que se extiende por encima de ese punto, hasta el límite superior de la atmósfera.

Todos los materiales y los seres vivos que poblamos este planeta estamos sometidos a la presión atmosférica. Si no somos conscientes de este peso del aire es porque ya estamos adaptados porque se ejerce por igual en todas direcciones y nuestros líquidos internos están a la misma presión.

El descubrimiento de la presión atmosférica

La fuerza del peso del aire no se descubre hasta 1643. Fue el científico italiano Evangelista Torricelli quien ese año hizo un experimento que demostró que el aire ejercía presión.

La imagen 1 muestra cómo lo hizo: en primer lugar, llenó con mercurio un tubo de un metro de longitud cerrado por un extremo y tapó el extremo abierto con el dedo. Luego, lo introdujo invertido en una cubeta llena de mercurio y, finalmente, retiró el dedo con cuidado para que no entrara aire en el tubo. En ese momento, el mercurio descendió hasta una altura de 760 mm sin llegar a vaciar el tubo y dejó un vacío en el extremo cerrado del tubo. Así fue como Torricelli también fue el primero en establecer cómo medir la presión atmosférica.

Imagen 1. Experimento de Torricelli

En la antigüedad, se creía que el vacío no era algo natural y que la naturaleza se resistía a tolerar la ausencia de aire. Torricelli demostró con su experimento que no es esa resistencia al vacío lo que impide al mercurio salir del tubo, sino que es la atmósfera la que ejerce presión sobre el mercurio de la cubeta. Por eso, la columna de mercurio desciende hasta una altura de 760 mm, el punto que iguala la presión que el aire ejerce sobre el mercurio de la cubeta.

b)

PL3_Pregunta - Mi aplicación x

No es seguro | lectura.uv.es/ReadAndLearn/PL3_Experimentos/PL3_Pregunta_Test?GrupoID=13588

Pregunta 5

¿La densidad del aire es igual en todos los puntos de la tierra?

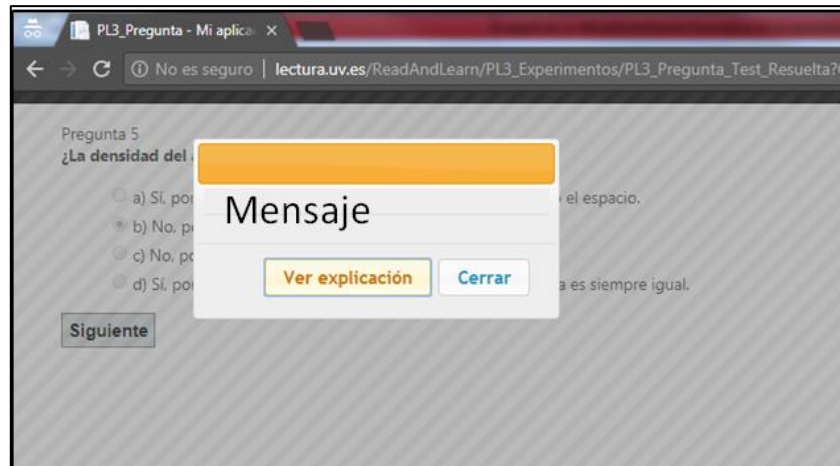
- a) Sí, porque los gases se expanden y tienden a ocupar todo el espacio.
- b) No, porque varía en función de la temperatura y la altura.
- c) No, porque el viento hace que varíe la densidad del aire.
- d) Sí, porque la mezcla de gases que componen la atmósfera es siempre igual.

Figura 1. Capturas de pantalla del texto (a), y de la pregunta (b)

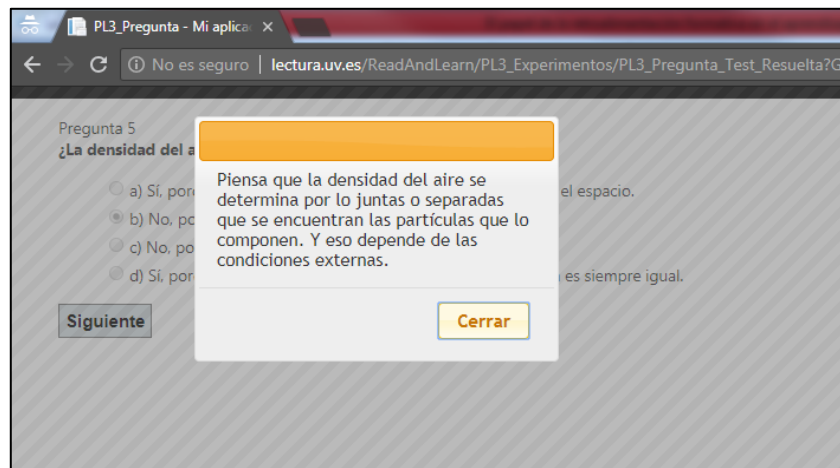
Tras responder cada pregunta, Read&Learn proporcionaba a los estudiantes mensajes de retroalimentación correctiva según la condición asignada, KR, KCR, o Control (Figura 2a). En la condición KR, los estudiantes recibían información acerca de la corrección de su respuesta (*Correcto* o *Incorrecto*), mientras que en la condición

KCR se informaba además acerca de cuál era la alternativa correcta (p.e., *Incorrecto. La respuesta correcta es: b) No, porque varía en función de la temperatura y la altura*). Por último, para la condición Control, los estudiantes recibieron mensajes neutros (*Has contestado la pregunta X*). Tras responder cada pregunta y recibir el mensaje de retroalimentación correspondiente a la condición, los estudiantes de las tres condiciones podían acceder voluntariamente al mensaje EF correspondiente con la explicación de la respuesta (Figura 2b). Al cerrar estos mensajes se presentó la pantalla de la pregunta con la alternativa seleccionada por el estudiante, pero sin señalar la respuesta correcta (Figura 2c). Nótese que sólo los estudiantes de la condición KCR conocían realmente la respuesta correcta cuando se equivocaban, ya que sólo esta retroalimentación incluía la respuesta correcta. Los estudiantes podían pasar a la siguiente pregunta clicando en el botón ‘Siguiente’ independientemente de si decidían leer EF o no.

a)



b)



c)

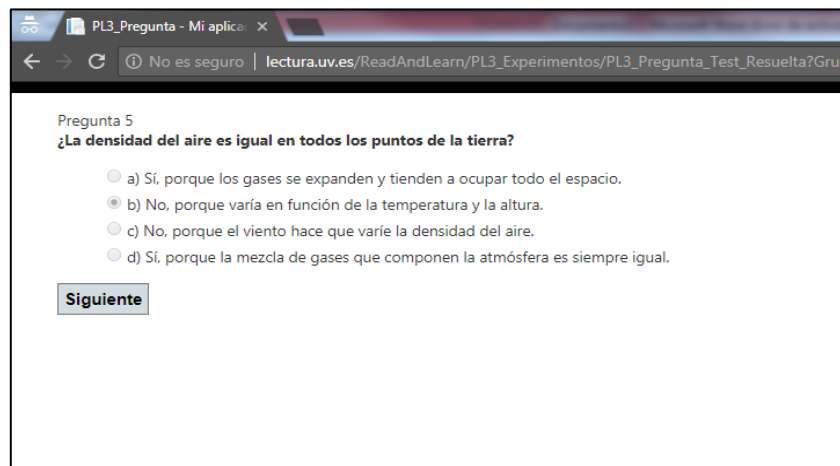


Figura 2. Capturas de pantalla del mensaje de retroalimentación correctiva recibido automáticamente después de responder (KR, KCR, o Control) (a), del mensaje EF al que podían acceder voluntariamente todos los estudiantes (b), y pantalla con la pregunta y sus alternativas tras cerrar los mensajes de retroalimentación (c)

2.4. Procedimiento

El investigador se puso en contacto con los orientadores de los centros educativos y les explicó el procedimiento del estudio. Después de conseguir los consentimientos informados, se concertaron las fechas en las que se aplicarían las pruebas. El procedimiento general del estudio puede verse en la Figura 3. La aplicación del estudio requirió de tres sesiones. En la primera sesión administramos la prueba de conocimiento previo en papel y lápiz en el aula ordinaria. En la segunda sesión aplicamos la tarea de aprendizaje en el aula de informática. Cada estudiante completó la tarea individualmente en un ordenador. Los estudiantes recibieron inicialmente instrucciones sobre el uso de Read&Learn (p.e., cómo pasar del texto a la pregunta, cómo contestar, y cómo acceder a los mensajes EF). Posteriormente, los estudiantes completaron la tarea a su ritmo, pudiendo buscar información en el texto al responder. Una vez respondida cada pregunta, Read&Learn proporcionó retroalimentación correctiva automáticamente (i.e., KR, KCR, o Control según la condición asignada). Posteriormente, todos los estudiantes tenían la opción de acceder a un mensaje EF para cada pregunta. Al cerrar los mensajes de retroalimentación, los estudiantes pudieron ver la pregunta con la opción seleccionada y el botón ‘Siguiente’ para continuar con la siguiente pregunta. El programa les avisó al finalizar la tarea. En la tercera sesión administramos la prueba final en papel y lápiz 24 horas después de completar la tarea de aprendizaje. Los estudiantes debían contestar las preguntas con la información que habían aprendido durante la sesión previa. Los estudiantes tenían un código asignado, conocido sólo por cada estudiante y su tutor, a fin de preservar su confidencialidad, a la vez que permitía identificar las respuestas de cada participante a lo largo de las tres sesiones.

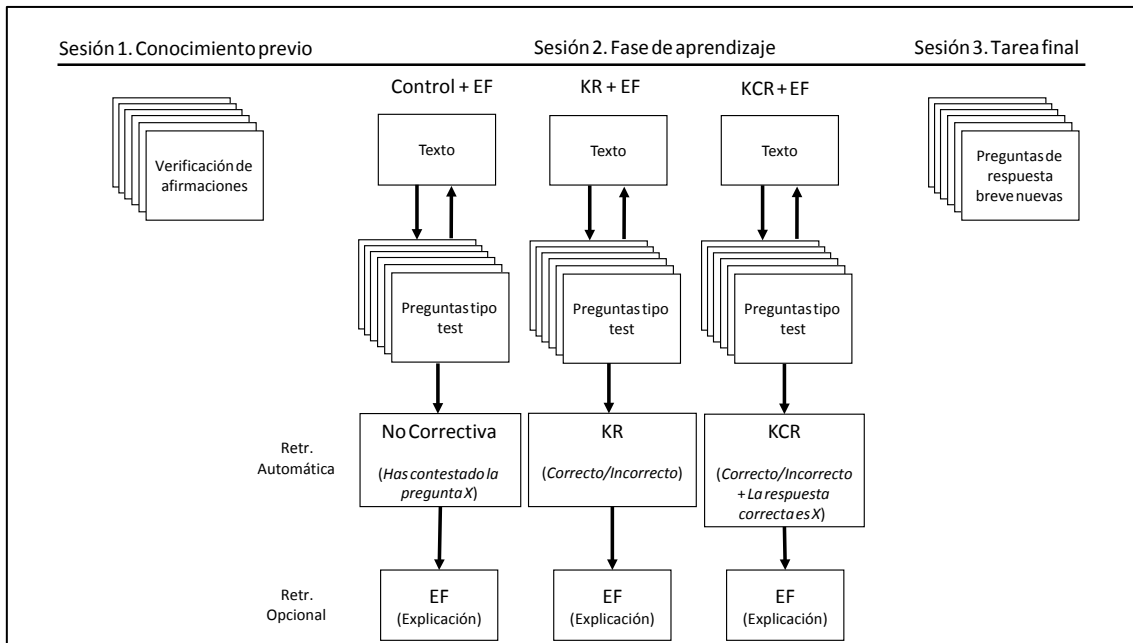


Figura 3. Esquema del procedimiento experimental utilizado para los grupos de retroalimentación Control, KR y KCR

2.5. Medidas

Se tomaron medidas para cada tarea. Respecto a la evaluación de conocimientos previos, calculamos el número de aciertos en la tarea de verificación de afirmaciones. Para la tarea de aprendizaje, Read&Learn proporcionó el rendimiento de los estudiantes como el porcentaje de respuestas correctas. Además, proporcionó dos medidas relacionadas con el uso de la EF: el porcentaje de estudiantes que decidieron acceder al menos a un mensaje y el porcentaje de preguntas en que los estudiantes decidieron acceder a la EF. Estas decisiones fueron condicionadas a la precisión de las respuestas de los estudiantes (correctas e incorrectas). Así calculamos los porcentajes de acceso a la EF en función de la exactitud de las respuestas, tomando como total la cantidad de accesos a los mensajes EF. Para la tarea final calculamos el porcentaje de respuestas correctas.

El estudio nos permitió comparar el acceso a la EF y el rendimiento en las tareas de aprendizaje y final en función del tipo de retroalimentación correctiva recibido. Para la medida de *acceso a la EF* utilizamos las pruebas no paramétricas Kruskal-Wallis y Wilcoxon, debido a que las puntuaciones no se distribuyen de forma normal. Los porcentajes de accesos a los mensajes EF en el grupo Control, $W(20) = 0.867$, $p = .01$, en el grupo KR, $W(24) = 0.854$, $p = .003$, y en el grupo KCR, $W(23) = 0.654$, $p < .001$,

se desvían significativamente de la distribución normal. No obstante, los análisis de homogeneidad de varianzas mediante la prueba de Levene demuestran que no existen diferencias significativas dependiendo del tipo de retroalimentación, $F(2, 64) = 1.26, p = .29$.

Utilizamos ANCOVAs para las medidas de producto, ya que las puntuaciones se ajustan a la distribución normal y las varianzas se distribuyen de forma homogénea. Para la *tarea de aprendizaje*, los aciertos en el grupo control, $W(20) = 0.953, p = .418$, y en el grupo KR, $W(24) = 0.938, p = .145$, tienen una distribución normal; no obstante, las puntuaciones del grupo KCR, $W(23) = 0.862, p = .005$, se desvían de la distribución normal. El análisis de homogeneidad de varianzas mediante la prueba de Levene demuestra que no existen diferencias significativas entre los grupos, $F(2, 64) = 2.83, p = .067$. Para la *tarea final*, los aciertos en el grupo control, $W(20) = 0.916, p = .082$, en el grupo KR, $W(24) = 0.976, p = .82$, y en el grupo KCR, $W(23) = 0.956, p = .384$, tienen una distribución normal. Además, la prueba de Levene demuestra que no existen diferencias significativas en las varianzas entre los tipos de retroalimentación, $F(2, 64) = 0.09, p = .92$.

3. RESULTADOS

3.1. Efecto del tipo de retroalimentación correctiva en el uso de la EF

Para analizar el efecto del tipo de retroalimentación correctiva (Control, KR, y KCR) en el porcentaje de estudiantes que utilizan la EF y el porcentaje de accesos a la EF utilizamos la prueba Kruskal-Wallis. Las medias y desviaciones típicas figuran en la Tabla 1. Para el porcentaje de estudiantes que acceden a la EF, encontramos un efecto significativo en función del tipo de retroalimentación correctiva recibido, $H(2) = 16.15, p < .001$. La prueba Jonckheere revela una tendencia significativa en los datos, indicando que a medida que la retroalimentación correctiva incluye información más específica sobre la respuesta (KR y KCR), el porcentaje de estudiantes que accede a la EF disminuye, $J = 746, z = -3.89, p < .001, r = -.48$. Para el porcentaje de accesos a la EF encontramos resultados paralelos, ya que éste también está significativamente influido por el tipo de retroalimentación correctiva recibido, $H(2) = 21.28, p < .001$. Además, la prueba Jonckheere revela una tendencia significativa, indicando que a medida que la retroalimentación correctiva incluye información más específica sobre la

respuesta (KR y KCR), el porcentaje de accesos a la EF disminuye drásticamente, $J = 746$, $z = -4.80$, $p < .001$, $r = -.59$. Ambos resultados confirman las hipótesis relacionadas con el efecto de la retroalimentación correctiva en el uso de la EF.

Tabla 1. *Estadísticos descriptivos de las decisiones de acceso a la EF en función del tipo de retroalimentación correctiva.*

	Control	KR	KCR
Estudiantes Acceden EF (%)	100 (0)	87.50 (33.78)	52.17 (51.08)
Accesos (%)	60.83 (32.12)	37.85 (29.49)	18.84 (29.86)

Nota. DT se muestra entre paréntesis.

Para comprender mejor el efecto de la retroalimentación correctiva en el uso de la EF analizamos el acceso a la EF en función de la precisión de las respuestas (Correctas, Incorrectas) para cada tipo de retroalimentación correctiva (Control, KR, y KCR) mediante la prueba Wilcoxon. Para el grupo Control, no existen diferencias en el porcentaje de accesos a los mensajes EF entre las respuestas incorrectas ($Mdn = 50.00$) y las correctas ($Mdn = 50.00$), $z = -1.42$, $p = .157$, $r = -.22$. Sin embargo, para el grupo KR encontramos que los estudiantes tienden a acceder a los mensajes EF más frecuentemente para las respuestas incorrectas ($Mdn = 77.50$) que para las respuestas correctas ($Mdn = 0$), $z = -3.75$, $p < .001$, $r = -.54$. Igualmente, para el grupo KCR, los estudiantes tienden a acceder a los mensajes EF más frecuentemente para las respuestas incorrectas ($Mdn = 16.67$) que para las respuestas correctas ($Mdn = 0$), $z = -2.02$, $p = .043$, $r = -.30$. Por tanto, los estudiantes del grupo Control accedieron a EF independientemente de la precisión de sus respuestas, mientras que los estudiantes que recibieron KR y KCR accedieron a EF después de responder incorrectamente (Figura 4). Por tanto, estos resultados confirman nuestra hipótesis inicial.

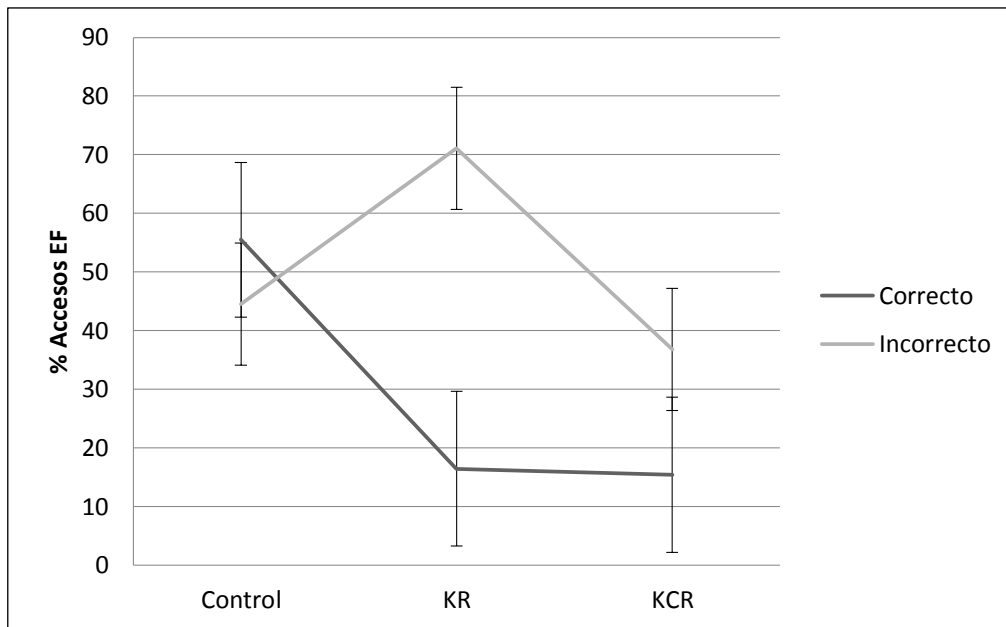


Figura 4. Porcentaje de accesos a la EF en función del tipo de retroalimentación correctiva y la precisión de las respuestas. El porcentaje ha sido calculado sobre el total de accesos de cada participante

3.2. Efecto del tipo de retroalimentación en el rendimiento de los estudiantes

Los porcentajes de acierto para la tarea de aprendizaje y la tarea final fueron analizados mediante dos ANCOVAs en los que introdujimos el tipo de retroalimentación correctiva (Control, KR, y KCR) como variable entre-sujetos, y el nivel de conocimiento previo como covariable. Las medias y desviaciones típicas para cada variable de rendimiento figuran en la Tabla 2. Los resultados del ANCOVA unifactorial para el porcentaje de aciertos en la tarea de aprendizaje no revelan diferencias significativas en función de la retroalimentación recibida, $F(2, 63) = 0.80, p > .05$. Sin embargo, observamos diferencias estadísticamente significativas en función del nivel de conocimiento previo de los estudiantes, $F(1, 63) = 7.51, p = .008, \eta_p^2 = .11$. Para el rendimiento en la tarea final encontramos resultados paralelos. No existen diferencias significativas en función de la retroalimentación recibida, $F(2, 63) = 1.45, p > .05$. Sin embargo, observamos diferencias estadísticamente significativas en función del nivel de conocimiento previo, $F(1, 63) = 30.60, p < .001, \eta_p^2 = .33$. Estos resultados confirman nuestra segunda predicción para el efecto de la retroalimentación en el rendimiento de los estudiantes.

Tabla 2. *Estadísticos descriptivos del rendimiento en las tareas de aprendizaje y final en función del tipo de retroalimentación.*

	Control	Verificación	Resp. Correcta
Aciertos Aprendizaje (%)	51.67 (16.80)	46.53 (10.97)	50.36 (19.38)
Aciertos Tarea Final (%)	39.25 (21.04)	38.33 (18.16)	43.91 (20.05)

Nota. DT se muestra entre paréntesis.

4. DISCUSIÓN

La retroalimentación formativa es uno de los componentes centrales en el diseño de entornos digitales de aprendizaje (Shute y Rahimi, 2017; van der Kleij et al., 2015). Aunque su efectividad ha sido ampliamente investigada (p.e., Hattie y Gan, 2011; Shute, 2008), existe poca investigación acerca de cuándo y cómo se utiliza la retroalimentación en entornos digitales (Fox et al., 2014; Máñez et al., 2019; Timmers y Veldkamp, 2011). Una complicación adicional es que frecuentemente no se compara la influencia que ejercen unos tipos de retroalimentación sobre otros, i.e., KR y KCR sobre EF. Por ello, el objetivo principal del estudio fue examinar el impacto de la retroalimentación correctiva (KR y KCR) en la decisión voluntaria de acceder a EF cuando los estudiantes aprenden conocimientos conceptuales de ciencias, una de las áreas en las que los estudiantes de Educación Secundaria experimentan más dificultades (OECD, 2006 y 2010). Un segundo objetivo fue examinar la influencia de la retroalimentación en el rendimiento final de los estudiantes. Estos objetivos fueron abordados con una de las tareas más representativas en el ámbito académico consistente en contestar preguntas a partir de un texto disponible (Ness, 2011). Así, este estudio contribuye a la comprensión del uso de la retroalimentación formativa y proporciona información valiosa para el diseño de entornos digitales de aprendizaje.

Respecto al primer objetivo, según predijimos, existe una marcada tendencia en el uso de la EF en función de la retroalimentación correctiva recibida. Proporcionar retroalimentación correctiva cada vez más detallada (KR y KCR) disuade a los estudiantes de acceder a los mensajes EF, seguramente porque muchos estudiantes se centran en conocer la corrección de sus respuestas, más que en un aprendizaje conceptual profundo (Máñez et al., 2019). Así, los estudiantes que desconocen si sus respuestas son correctas tienden a acceder a un mayor número de mensajes EF, probablemente con la intención de encontrar pistas que les permitan confirmar su

modelo de respuesta o conocer cuál es la respuesta correcta. Los estudiantes que reciben KR acceden a los mensajes EF en menor medida que los de control ya que tienden a acceder cuando fallan, pero no cuando aciertan. Los estudiantes que reciben KCR acceden a los mensajes EF en muy pocas ocasiones ya que cuando fallan reciben información sobre cuál es la respuesta correcta. De hecho, aproximadamente la mitad de los estudiantes que recibió retroalimentación KCR nunca decidieron consultar la EF. Parece que invertir esfuerzos en procesar retroalimentación adicional cuando se conoce la respuesta correcta (KCR) se considera poco necesario y, por tanto, no se implican en el coste que conlleva leer las explicaciones de la EF. Es llamativo que todos los estudiantes del grupo control decidieron acceder al menos a un mensaje EF, siendo este porcentaje significativamente más bajo en las otras condiciones, particularmente en la condición KCR donde casi el 50% de los estudiantes nunca accedió a un mensaje EF. Aunque resultados previos indican que los estudiantes acceden frecuentemente a la EF en entornos digitales (Timmers y Veldkamp, 2011), nuestros resultados sugieren que los estudiantes deciden utilizar la EF en un número limitado de ocasiones. Cabe destacar que Timmers y Veldkamp (2011) diseñaron mensajes EF que incluían la respuesta correcta, por lo que la decisión de utilizar esa retroalimentación pudo estar motivada por la retroalimentación KCR incluida. En nuestro estudio, los resultados sugieren que los estudiantes acceden a la EF para intentar conocer la respuesta correcta, ya que la tasa de accesos depende de la información correctiva recibida automáticamente tras responder.

Por tanto, ofrecer retroalimentación correctiva interfiere en cierta medida en la decisión de buscar EF adicional. ¿Cuál es, pues, la función de la retroalimentación correctiva? Los resultados sugieren que la función de la retroalimentación KR y KCR es verificar el modelo de respuesta creado por el estudiante para responder y mejorar la monitorización del uso de la EF (Mory, 2004). Así, cuando los estudiantes reciben retroalimentación de tipo correctivo y fallan, entonces deciden acceder a EF (Fox et al., 2014; Timmers y Veldkamp, 2011). Cuando los estudiantes no reciben retroalimentación correctiva acceden a la EF independientemente de si sus respuestas son correctas o incorrectas, lo que resulta congruente con resultados previos que indican que los estudiantes experimentan problemas al monitorizar la precisión de sus respuestas (Dunlosky y Metcalfe, 2009). Aunque el uso de la EF parece estar motivado por la precisión de las respuestas de los estudiantes, hay alumnos que, a pesar de ser informados de sus errores, deciden no acceder a la EF. Puede que estos estudiantes consideren la tarea demasiado compleja, no estén suficientemente motivados para

buscar información adicional, o tengan dificultades para entender la EF. Tomados en conjunto, los resultados de nuestro estudio son consistentes con la idea de que el objetivo de los estudiantes al responder preguntas de un texto disponible es proporcionar la respuesta correcta (Rouet et al., 2017).

Respecto al segundo objetivo, examinar el efecto del tipo de retroalimentación en el rendimiento de los estudiantes, los resultados revelaron que no existen diferencias entre los tres tipos de retroalimentación empleados (Control+EF, KR+EF, y KCR+EF), resultado consistente con investigaciones recientes en las áreas de comprensión y aprendizaje de ciencias (Golke et al., 2015; Maier et al., 2016). Este resultado puede deberse al uso de la EF que realizaron los estudiantes en función de la retroalimentación correctiva recibida automáticamente después de responder cada pregunta. Como predijimos, el mayor uso de la EF por parte de los estudiantes de los grupos Control y KR pudo haber compensado el efecto positivo de conocer la respuesta correcta en la condición KCR. En apoyo de esta interpretación se puede mencionar el dato que los estudiantes del grupo Control no sólo tuvieron acceso a la EF, sino que además hicieron un mayor uso de esta información. Los estudiantes que recibieron retroalimentación KCR, por su parte, decidieron acceder a menos mensajes EF. Sin embargo, el hecho de conocer la respuesta correcta pudo tener un efecto positivo en su rendimiento (van der Kleij et al., 2015), ya que conocer la respuesta correcta permite modificar el modelo de respuesta construido durante la contestación a la pregunta. Así, es posible que la retroalimentación KCR sea capaz de influir en el aprendizaje profundo de este tipo de conocimientos de ciencias. Parece que proporcionar retroalimentación que incluya información correctiva y elaborada puede ejercer una influencia positiva en el aprendizaje de los estudiantes cuando contestan preguntas de comprensión en entornos digitales (Llorens et al., 2016; Llorens et al., 2015). Aun así, debemos ser cautos con respecto a la efectividad de la retroalimentación en este tipo de tareas, ya que los porcentajes de aciertos fueron relativamente bajos, debido probablemente a la dificultad de los conocimientos del texto.

Estos hallazgos tienen implicaciones teóricas y prácticas destacables. Este experimento aporta información teórica acerca de cómo los estudiantes de Secundaria monitorizan y autorregulan el uso de la retroalimentación cuando han de contestar preguntas de un texto de ciencias. Los resultados de este estudio ayudan a conocer la tendencia de los estudiantes a utilizar voluntariamente la EF en función de la presencia de retroalimentación KR y KCR. Igualmente, ofrece información valiosa para el diseño

de herramientas digitales en las que se proporcione retroalimentación al responder preguntas sobre los contenidos de un texto. Aunque la retroalimentación correctiva KR y KCR pueden interferir en el uso de la EF, parece necesario que los entornos digitales proporcionen ambos tipos de retroalimentación (Lefevre y Cox, 2016; Mason y Bruning, 2001), especialmente cuando se promueve el aprendizaje de conocimientos complejos (Tärning, 2018). Aunque nuestro estudio sugiere que el uso de la EF es reducido, especialmente cuando se recibe retroalimentación KCR, esta información parece ser útil para corregir el modelo de respuesta creado por los estudiantes cuando se aprenden conceptos complejos en el área de ciencias.

El presente estudio tiene una serie de limitaciones que abren nuevas líneas de estudio para explorar el uso y efectividad de la retroalimentación formativa en entornos digitales. En primer lugar, la fase de aprendizaje tuvo una duración breve y los materiales evaluaron únicamente una temática del área de física. Aunque los patrones observados en la decisión de buscar EF después de recibir retroalimentación correctiva KR y KCR son claros, es necesario conocer si estos resultados se replican en otros dominios de conocimiento del aprendizaje de ciencias, así como en otras áreas como matemáticas o ciencias sociales. Asimismo, es importante señalar que los resultados muestran una marcada tendencia en el uso de la EF a pesar de contar con una muestra relativamente pequeña. Nuevos estudios son necesarios para replicar estos resultados con muestras más amplias que abarquen también otras etapas educativas. En segundo lugar, parece necesario incluir una condición control en la que los estudiantes no reciban ningún tipo de retroalimentación, ya que las diferencias en rendimiento respecto a los otros grupos experimentales pueden haberse minimizado debido a que la condición control accedió a los mensajes EF frecuentemente. En tercer lugar, el rendimiento en la prueba final fue relativamente bajo, por lo que parece necesario incluir una prueba tipo test, ya que los estudiantes no habían sido entrenados para responder preguntas en formato abierto. En cuarto lugar, sería interesante conocer el efecto de factores externos en el uso de la EF (Lefevre y Cox, 2016), ya que estos factores pueden ser fácilmente manipulables en los entornos digitales y pueden desempeñar un papel importante en el proceso de aprendizaje de los estudiantes. Por ejemplo, la disponibilidad de los textos al contestar las preguntas es una variable situacional que puede afectar al uso de la EF. Finalmente, resulta interesante conocer la razón por la que los estudiantes deciden no buscar EF. Así, examinar aspectos motivacionales puede ofrecer información valiosa sobre la autorregulación del uso de la EF (e.g., Nicol y Macfarlane-Dick, 2006), ya que

la motivación juega un papel central en situaciones de aprendizaje con retroalimentación (Kluger y DeNisi, 1996).

Tomados en conjunto, los resultados sugieren que no podemos asumir que la retroalimentación sea procesada automáticamente. Más bien, parece que los alumnos monitorizan el uso de la retroalimentación elaborada (EF) en situaciones de aprendizaje autorregulado (Butler y Winne, 1995), ajustando su uso en función de la retroalimentación correctiva KR o KCR proporcionada tras contestar cada pregunta y la precisión de sus respuestas. De este modo, los hallazgos de este experimento son valiosos para el diseño de entornos digitales de aprendizaje en los que la retroalimentación puede ajustarse a las necesidades de los estudiantes cuando aprenden conocimientos conceptuales.

Referencias

- Anmarkrud, Ø., McCrudden, M.T., Bråten, I., y 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., y Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213–238. doi:10.3102/00346543061002213
- Butler, D. L., y Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. doi:10.3102/00346543065003245
- Cerdán, R., Gilabert, R., y 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
- Dunlosky, J., y Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage.
- Fox, J.-P., Klein Entink, R., y Timmers, C. (2014). The joint multivariate modeling of multiple mixed response sources: Relating student performances with feedback behavior. *Multivariate Behavioral Research*, 49, 54-66. doi:10.1080/00273171.2013.843441
- Golke, S., Dörfler, T., y 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
- Hattie, J., y Gan, M. (2011). Instruction based on feedback. En R. E. Mayer y P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249 – 271). New York, NY: Routledge.

- Kluger, A. N., y 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
- Lefevre, D., y Cox, B. (2016). Feedback in technology-based instruction: Learner preferences. *British Journal of Educational Technology*, 47(2), 248-256. doi:10.1111/bjet.12227
- Llorens, A. C., Vidal-Abarca, E., y 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., y 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
- Maier, U., Wolf, N., y Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85-98. doi:10.1016/j.compedu.2015.12.002
- Máñez, I., Vidal-Abarca, E., Kendeou, P., & Martínez, T. (2019). How do students process complex formative feedback in question-answering tasks? A think-aloud study. *Metacognition and Learning*. doi:10.1007/s11409-019-09192-w
- Mason, B. J., y Bruning, R. (2001). *Providing feedback in computer-based instruction: What the research tells us*. Center for Instructional Innovation, University of Nebraska–Lincoln. Recuperado el 20 de noviembre de 2017 de

https://www.researchgate.net/publication/247291218_Providing_Feedback_in_Computer-based_Instruction_What_the_Research_Tells_Us

Mory, E. H. (2004). Feedback research revisited. En D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Erlbaum.

Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. *Digital Education Review*, 23, 7-26.

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Gogvadze, G., y Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, 71, 56-76. doi:10.1016/j.compedu.2013.09.011

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

Nicol, D. J., y Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi: 10.1080/03075070600572090

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

OECD (2010). *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*. Paris: OECD.

Roelle, J., y Berthold, K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, 49(2), 142-156, doi:10.1016/j.learninstruc.2017.01.008

- Rouet, J-F., Britt, M. A., y Durik, A. M., (2017). RESOLV: Readers' Representation of Reading Contexts and Tasks. *Educational Psychologist*, 52(3), 200-215. doi:10.1080/00461520.2017.1329015
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. doi:10.3102/0034654307313795
- Shute, V. J., y Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33(1), 1-19. doi:10.1111/jcal.12172
- Tärning, B. (2018). Review of feedback in digital applications-Does the feedback they provide support learning? *Journal of Information Technology Education: Research*, 17, 247-283. doi:10.28945/4104
- Timmers, C. F., y 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., Eggen, T. J. H. M., Timmers, C. F., y Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, 58, 263–272. doi:10.1016/j.compedu.2011.07.020
- Van der Kleij, F. M., Feskens, R. C., y 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., Mañá, A., y 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

CHAPTER 5

STUDY 4: DO SECONDARY-SCHOOL STUDENTS USE ELABORATED FEEDBACK? THE ROLE OF KNOWLEDGE-OF-RESPONSE FEEDBACK, TEXT AVAILABILITY AND PRIOR KNOWLEDGE

ABSTRACT

This study examines how Knowledge-of-Response feedback (KR) and text availability influence Secondary-school student's decision to use Elaborated Feedback (EF) when answering multiple-choice questions from a science text. Half the participants had the text available when answering the questions and half did not. Half the participants received item-based KR feedback and half non-corrective feedback. All the participants were allowed to access item-based EF with an explanation on the knowledge assessed. A final test was administered 24 h later. Results showed that keeping the text unavailable increased EF use and that KR feedback increased EF use after failure. Further, EF moderated the relation between student's prior knowledge and performance. However, both prior knowledge and EF played independent roles when predicting final performance.

Keywords: Elaborated Feedback (EF); Knowledge-of-Response feedback (KR); text availability; prior knowledge; learning; science

1. INTRODUCTION

Answering questions from a text is a common activity for assessment and learning purposes in school settings (Ness, 2011; OECD, 2010). As happens in many academic subjects, students usually perform question-answering tasks to learn science knowledge. In such scenarios, formative feedback is a common and sometimes effective tool to overcome comprehension and learning difficulties. Feedback embraces a wide variety of feedback types (Shute, 2008). A common distinction is usually made between Knowledge of Response or KR (i.e., response verification feedback: *Correct/Incorrect*), Knowledge of Correct Response or KCR (i.e., information indicating the correct answer: *The correct response is X*), and Elaborated Feedback or EF (i.e., any further information beyond student's performance such as explanations, hints, or examples). The extant literature suggests that EF is a powerful source of information to enhance learning (Van der Kleij, Feskens, & Eggen, 2015) and that students take control of the EF use in digital environments (e.g., Máñez, Vidal-Abarca, Kendeou, & Martínez, 2019; Timmers & Veldkamp, 2011). Although it is common to assume that students process EF when received, one may wonder what external and internal factors may influence the students' decision to use EF in digital environments (Lefevre & Cox, 2016). The present study examines whether two external sources of information (i.e., KR feedback and text availability) influence Secondary-school student's decision to use EF made of explanations when learning conceptual knowledge of physics. Given that this is one of the science areas students usually struggle with, we took into account their prior knowledge.

Nowadays it is increasingly common to perform learning tasks in digital environments, since these systems are able not only to grade responses and deliver feedback automatically, but also to display templates in which contextual variables can be modified easily. For instance, it is possible to deliver KR feedback after answering each question and give students the opportunity to access additional EF at their will. Likewise, it is also possible to design an environment in which texts are (un)available to the students while solving learning tasks. Recent research suggests that students pay attention to KR and KCR feedback in digital environments (Lefevre & Cox, 2016; Máñez et al., 2019), and that having the text available may play a part in learning (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008). However, to the best of our

knowledge, no study has examined how these external sources of information influence the student's decision to use EF in a digital environment.

1.1. Feedback effectiveness and use

Feedback is a common learning procedure aimed at increasing students' knowledge or learning skills (Shute, 2008). Although there is a widespread perception about the positive effects of feedback on learning outcomes, there is wide variability in findings among research studies (e.g., Kluger & DeNisi, 1996; Mory, 2004; Shute, 2008; Stobart, 2008). Whereas some studies have reported positive effects, others have found no effects or even negative effects on learning (for a review, see Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Van der Kleij et al., 2015). This variability likely is the result of differences in tasks, participants, and feedback characteristics (Shute, 2008). As we have mentioned above, there are different types of feedback based on their complexity. Despite the fact that EF embraces different types of information, it is the most effective for learning purposes (e.g., Jaehnig & Miller, 2007; Mory, 2004; Shute, 2008; Van der Kleij et al., 2015). In a recent meta-analysis, Van der Kleij et al. (2015) reported that larger effect sizes are related to EF messages in comparison to KCR and KR feedback messages. Further, results showed that EF plays an essential role in higher order learning outcomes, probably because this information allows learners to bridge the gap between their actual and the desired level of knowledge. For instance, Butler, Godbole, and Marsh (2013) provided evidence that EF made of explanations was more beneficial than KCR to transfer knowledge.

However, EF is not effective per se. In the area of learning science knowledge in a digital environment, Maier, Wolf, and Randler (2016) conducted an experiment to examine whether three different types of feedback influenced performance. Secondary-school students studied and answered questions about a biology text under three feedback conditions: EF (i.e., KR feedback and a short explanation after incorrectly answered questions), KR feedback, and no feedback (i.e., reading appropriate texts related to the biology topics). Students in the EF condition also self-reported their use of EF. Initial results showed that there were no significant differences among conditions. However, when researchers divided the students in the EF condition into those who used and did not use EF according to the students' self-reported questionnaire, they found that learners who received and used EF and those who received KR feedback

outperformed students in the no-feedback condition and those who received EF but did not use it. This finding suggests that making use of EF may play a role in learning science knowledge.

Therefore, it seems that students have to be willing to process the feedback content actively to be effective (Bangert-Drowns et al., 1991). Recent literature has shown that students tend to compare their performance to the correct response when they are allowed to self-regulate their learning within a testing and feedback learning environment (Dunlosky & Rawson, 2015). However, little research has examined the extent to which young learners engage in feedback processing in digital environments (Van der Kleij, Eggen, Timmers, & Veldkamp, 2012). Recent evidence suggests that learners focus primarily on knowing the correctness of their answers or the correct response rather than processing other elaborative feedback components (Lefevre & Cox, 2016; Máñez et al., 2019). Further, previous evidence suggests that providing learners with KCR feedback reduces the likelihood of using additional EF made of explanations (Máñez, submitted for publication).

In a similar vein, Timmers and Veldkamp (2011) and Fox, Klein Entink, and Timmers (2014) investigated the students' use of EF messages provided after completing a computer-based test on information literacy. Students completed the assessment and received an overview of their correct and incorrect answers (i.e., KR feedback) and had the option to access item-based EF, which consisted of KCR feedback along with an explanation. Main findings showed that students accessed EF messages more often after incorrect as opposed to correct answers. However, these EF messages included KCR feedback, so that students may access EF to knowing the correct response rather than processing the explanations. Our study overcame that limitation by providing optional EF made of explanations aimed at promoting understanding of conceptual knowledge in the area of physics, but without stating the correct response. In our experiment, Secondary-school students answered questions from a science text about the atmospheric pressure and the wind phenomenon, received either KR or control feedback after responding each question and had the option to access EF voluntarily. For example, for the question '*If Torricelli would have used a 2-meter tube full of mercury, instead of a one-meter tube, and would have repeated his experiment, what would have happened to the mercury inside the tube?*' (Correct response: '*It would have stopped at the same level, 760mm*'), the corresponding EF message was: '*Think that the level at which the mercury stops is determined by the point*

where the weight of the mercury inside the tube and the weight of the air column above the bucket are equal'. This EF message contains an explanation to facilitate the understanding of Torricelli's experiment to demonstrate atmospheric pressure, although without providing the correct answer.

Further research is necessary to examine whether and how KR feedback influences the EF use compared to a control condition in which no corrective feedback is delivered. Since KR feedback serves to verify responses, its delivery may influence the way students decide to use EF (i.e., students may access EF more frequently after providing incorrect responses), whereas receiving no corrective feedback may make students use EF for both correctly- and incorrectly-answered questions (Máñez, submitted for publication). Likewise, EF use may also be influenced by other contextual factors such as having the learning materials (un)available while performing the task.

1.2. Text availability

Among the external factors that may influence how students use EF (e.g., type of corrective feedback), text availability may also play a part. Previous evidence suggests that having the learning materials available influences positively the student's learning outcomes, although this benefit disappears after a delay (Agarwal et al., 2008). Ozuru, Best, Bell, Witherspoon, and McNamara (2007) had undergraduate students answer questions from an expository science text either with or without the text available during the question-answering task. Results showed that having access to the text increased students' performance. Recent findings suggest that answering questions with the text available increases performance on open-ended questions, but has little or no effect on multiple-choice questions (Ferrer, Vidal-Abarca, Serrano, & Gilabert, 2017). Further, text availability also influences how students use the text information (Ferrer et al., 2017; Higgs, Magliano, Vidal-Abarca, Martínez, & McNamara, 2017), so that having this source of information at hand or not may have an impact on students' decision to use EF. Students can search the text for question-relevant information under available-text assessments, whereas they have to answer the questions based on the mental representation built when studying the text initially under unavailable-text assessments (Ferrer et al., 2017). Thus, students who have the text unavailable may use other sources of information at hand (e.g., EF messages) more often than those students

who have the text available. This feedback-seeking behavior seems to be the only way to modify their mental representation as they go through the task.

Moreover, student's decision to use EF may vary as a function of both KR feedback and text availability. Imagine a learner answering questions without having access to the text and receiving no corrective feedback on their responses, but having access to EF. In such case, the students may be more prone to access EF in an attempt to uncover the correct response. On the contrary, imagine a student who have the text available while answering the questions and receiving item-based KR feedback. Those students may use EF less frequently since they may consider unnecessary any additional information. The study of this feedback-seeking behavior as a function of these external factors (i.e., KR feedback and text availability) provides valuable information for the design of digital learning environments.

1.3. Prior knowledge

Among the internal factors that may influence feedback use and effectiveness, learner's prior knowledge is arguably one of the most relevant individual variables when it comes to reading comprehension (e.g., Kendeou & O'Brien, 2018; McNamara & Kintsch, 1996) and learning science knowledge from texts (e.g., Bohn-Gettler & Kendeou, 2014; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Mason, Tornatora, & Pluchino, 2013). Regarding feedback, it is largely unknown the extent to which students' prior knowledge may interact with feedback use when learning conceptual knowledge from texts. Previous research suggests that students' prior knowledge may moderate the effects of different types of feedback (Smits, Boon, Sluijsmans, & Van Gog, 2008). Empirical evidence in the field of mathematics has shown that both elementary-school and university students with lower prior knowledge tend to benefit more from feedback than students with higher prior knowledge (e.g., Fyfe, Rittle-Johnson, & DeCaro, 2012; Krause, Stark, & Mandl, 2009). Therefore, those students with low prior knowledge who engage in using EF frequently may benefit from it especially.

1.4. The current study

We had two aims in this study. The first aim was to examine the extent to which KR feedback and text availability influenced Secondary-school students' decision to

access additional item-based EF, while taking into account the student's prior knowledge. We hypothesized that unavailable-text assessment would increase EF access because the lack of text information would enhance student's willingness to use additional EF (Hypothesis 1). We also hypothesized that KR feedback would trigger lower rates of EF accesses compared to the control non-corrective feedback because students who receive KR feedback would use EF mainly after failure (Hypothesis 2). Further, we anticipated an interaction effect between the KR feedback and the text availability factors (Hypothesis 3). Thus, students who answered the questions under the unavailable-text condition and non-corrective feedback may be more prone to access EF messages since no external sources of information are available during the question-answering process.

The second aim was to determine the extent to which KR feedback and text availability influenced the student's learning outcomes. We hypothesized that having the text available and receiving KR feedback would enhance student's performance during the learning task, although this positive effect may disappear after a delay (Hypothesis 4). An additional hypothesis was formulated about how EF access may affect the relation between the learner's prior knowledge and their learning outcomes (Hypothesis 5). It may be the case that EF does not affect this relation, hence EF and prior knowledge would have independent effects on learning. Alternatively, it may be that EF moderates the relation between prior knowledge and learning, hence the EF use would reduce the positive effects of prior knowledge.

2. METHOD

2.1. Participants and Design

A total of one hundred fifty-seven 9th graders aged 14-16 years ($M = 14.54$, $SD = .58$) participated in this study (55.7% female). All participants were Spanish native speakers. The data of eighteen students were excluded from data analysis due to absences during one or more phases of the study ($n = 10$), lack of parents' consent to participate in the study ($n = 5$), software failures ($n = 1$), grade repetition ($n = 1$), or inability to complete the task ($n = 1$). Participation was approved by the school's board. The study was approved by the Research Ethics committee of the University of Valencia and followed the guidelines of the Declaration of Helsinki.

After assessing students' prior knowledge in Science, participants were assigned to one of four experimental conditions. The experiment constituted a 2 (text availability: unavailable vs. available) by 2 (KR feedback: KR vs. control) between-subjects design. There were 34 students in the unavailable text - control feedback group, 36 students in the unavailable text - KR feedback group, 34 students in the available text - control feedback group, and 35 students in the available text - KR feedback group. All students had studied Natural Sciences and Physics according to the Spanish academic curriculum. Thus, students were expected to have sufficient prior knowledge to comprehend the topic of the reading task: Atmospheric pressure and wind phenomenon.

2.2. Materials

Materials included a prior knowledge test, the learning assignment task, and a final test. Prior knowledge in Science was measured with a 30-item sentence verification test with three choices: 'True/False/I do not know'. The test included general (*The temperature is a property of the matter.*) and topic items (*When the temperature of a gas decreases, its pressure decreases.*). Each correct response was scored 1 and each incorrect a zero; 'I do not know' responses were counted separately.

For the learning assignment task, we developed a science text, 12 multiple-choice comprehension questions and the corresponding 12 EF messages. The science text focused on the atmospheric pressure and the wind phenomenon. The text was created drawing on Secondary-school science textbooks and was not previously taught to the students. The 1126 words length text was divided into 30 segments of one or several sentences each, including (sub)heading(s) and figure notes. Regarding readability indexes, the Flesch-Szigriszt reading grade level was 70.91, indicating that the text can be considered "Quite easy" to read according to the INFLESZ Scale (Barrio-Cantalejo et al., 2008). We developed 12 multiple-choice questions, most of which required students to focus on the gist meaning of text sentences and make text-based inferences. For the question '*Is the density of the air equal at all places on the Earth?*', the correct answer is '*No, because it varies depending on temperature and altitude*'. The text included three sentences explaining that the atmosphere is an extremely huge space made of particles whose distribution is not uniform (sentence 1), that the density of the layers of the air varies depending on the altitude (the size of the layers on the top is different than on lower points) (sentence 2), and atmospheric

pressure varies also depending on the temperature (sentence 3). Other questions, however, asked students to solve problems by applying text information to new situations. For the question *‘If you try to replicate the experiment of Torricelli on top of an eight thousand meter mountain rather than at sea level, where Torricelli conducted the original one, what do you think will happen?’*, the correct answer is *‘More mercury of the pipe will come out in the bucket, descending more than 760mm’*. These questions involved more text information. For example, relevant sentences to answer this question included the following information: when Torricelli conducted his experiment, the mercury descended to 760 mm; Torricelli demonstrated that the atmosphere put pressure on the mercury of the bucket; the mercury come out up to 760mm because at this point the pressure of the mercury of the pipe equates the pressure the air exerts on the bucket; the pressure of the air at the sea level is equal to the pressure of a column of mercury (Hg) 760mm high, a value equal to 1 atmosphere ($760\text{mmHg} = 1 \text{ atm}$); the density of the air varies depending on the altitude, since the size of the layers of air varies depending on the altitude; lower layers of the atmosphere support the weight of layers on top, which increases the air density; the pressure on top of a mountain of 3000 meters is lower than on the beach. Item-based EF messages were developed. EF began with the expression *‘Think that...’* followed by an explanation to help students understand the knowledge assessed, but without stating the correct answer. For instance, the EF for the last question was *‘Think that as you are at higher altitude there is less air above; therefore, the weight of air and mercury inside the tube will be matched immediately after it has come out a different quantity of mercury rather than when the experiment is done at sea level’*. Each correct response received a score of 1 and each incorrect a zero. Thus, the learning task is expected to help students acquire knowledge by means of deploying high-order comprehension processes.

The final test consisted of eight multiple-choice questions similar to those included in the learning assignment task. For example, the students answered questions such as *‘How is the density of the atmospheric air on top of a mountain?’* (correct answer: *‘It is lower because the column of air is smaller’*), or *‘Imagine that Torricelli would have used liquid gold in his experiment, which is more dense than mercury. What do you think it would have happened?’* (correct answer: *‘The liquid gold would have descended more than 760mm’*). Each correct response received a score of 1 and each incorrect response a zero.

2.3. Apparatus

Participants completed the learning assignment on Read&Learn, a computer-based system that records students' interactive behaviors. Read&Learn is the new and sophisticated version of Read&Answer (Vidal-Abarca et al., 2011), a program that traces online students' behaviors and processing times while performing task-oriented reading activities. Read&Learn also provides students with different feedback messages according to the students' performance. The Read&Learn template we used displays the text and the questions on different screens. Thus, the text was displayed on a screen and the question wording and its answer choices on another screen (Figure 1).

a)



b)

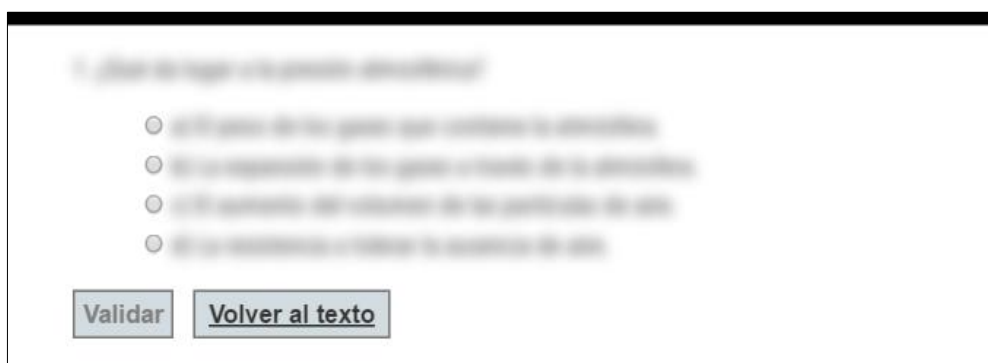


Figure 1. Screenshots of the text (a) and the question (b) screens with masked information. Please note that students under unavailable text condition did not have the “Back to text” button in the question screen

Text and multiple-choice question content were displayed using a masking procedure. To read text information, students had to click on each text segment to access its content. Once participants unmasked a piece of text information, the other segments remained masked until students clicked on them (Figure 2a). The same process applied for the question wording and answer choices (Figure 2b). The masking procedure allowed precision in recording students' information access patterns.

a)

The screenshot shows a web page with a title bar 'Preguntas' and a main heading 'La presión atmosférica y el vacío'. The page contains several paragraphs of text, each with a small grey box over it, indicating that the content is masked and can only be accessed by clicking on that specific area. The text includes definitions of atmospheric pressure and the concept of a vacuum. At the bottom, there is an illustration of a hand holding a glass tube vertically, dipping it into a container of liquid. The tube is partially submerged, and the liquid level inside the tube is higher than the level in the container. A label '760 mm' is placed next to the tube, indicating the height of the liquid column. The entire page is framed by a thin black border.

This screenshot is identical to the one above, showing the same web page with masked text segments and the illustration of the glass tube experiment. The layout, text, and graphics are the same, demonstrating the masking procedure used in the study.

b)

1. ¿Qué da lugar a la presión atmosférica?

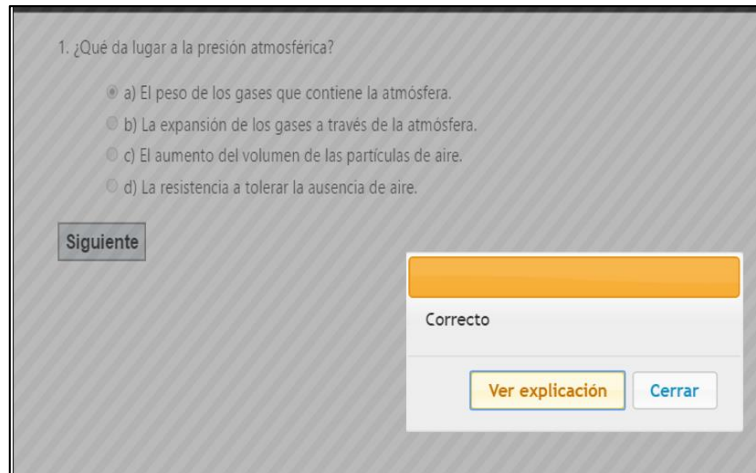
- a) El peso de los gases que contiene la atmósfera.
- b) La expansión de los gases a través de la atmósfera.
- c) El aumento del volumen de las partículas de aire.
- d) La resistencia a tolerar la ausencia de aire.

Validar

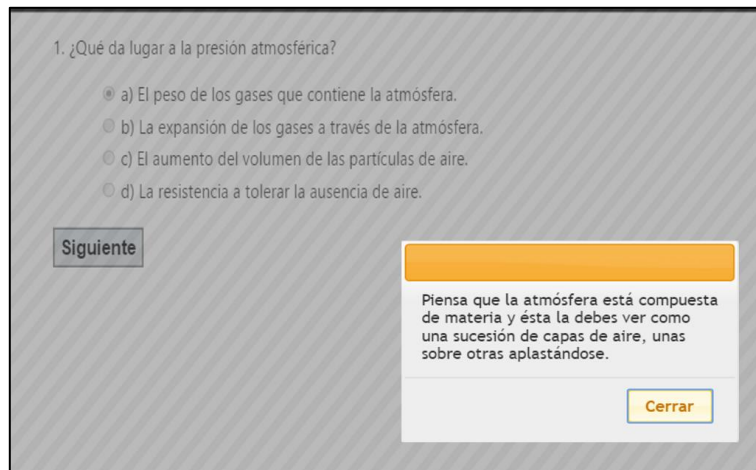
Figure 2. Masking procedure to read text (a) and question (b) content

After answering each question, Read&Learn provided students with feedback messages. For the KR feedback condition, students received *Correct* or *Incorrect* feedback messages depending on their answer correctness, whereas for the control feedback condition, students received neutral messages (e.g., “*You have answered question number 1*”) (Figure 3a). Following this feedback, all the participants had the option to click and access an EF message providing an explanation of the question answer (Figure 3b). These EF messages began with the statement “*Think that...*” followed by a brief explanation of the atmospheric phenomenon that the question assessed. After receiving the feedback, question content was presented unmasked with the students’ answer choice selected (Figure 3c), although the correct choice was not marked. Participants were able to move to the next question by clicking on the “Next” button on the screen.

a)



b)



c)

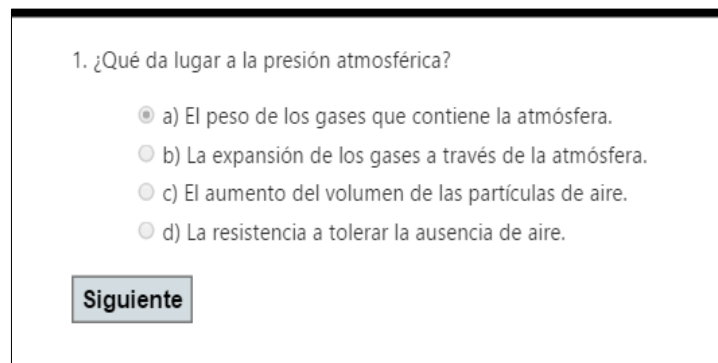


Figure 3. Screenshots of feedback depending on the KR feedback condition (a), the EF message (b), and the unmasked question screen with the students' answer choice selected (c)

The interface was slightly different depending on text availability. In the available text condition, students had access to a “Back to text” button in the lower part

of the question screen (Figure 1b) to return to the text screen during the question-answering process, and the “Questions” button on the text screen (Figure 1a) to return back to the question screen. This feature allowed us to distinguish clearly whether participants processed the text or the question content. In the unavailable text condition, students did not have access to those buttons/options.

2.4. Procedure

The experiment involved three sessions illustrated in Figure 4. In the first session, a sentence verification test was administered in paper-and-pencil format to assess the students’ prior knowledge in the regular classroom. Then, students received general instructions about the Read&Learn software (e.g., how to interact with the program, or how to access the EF messages). After scoring the sentence verification test, participants were equally distributed among the experimental groups based on their prior knowledge.

In the second session, the learning assignment task was administered in the computers room. Each student sat in front of a computer screen. Students were first reminded of the general software instructions and instructed that they would read a science text and answer a set of questions. Approximately half of the participants were told that they would not have the text available during the question-answering task and the other half were told that they would have the text available while answering the questions. Likewise, half of the participants were told that they would receive an automatic message on the question answered (Control feedback group) and the other half were told that they would receive automatic feedback on the correctness of their answers (KR feedback group). Moreover, all the participants were told that they would have the opportunity to access an additional item-based explanation (EF message) to better understand the knowledge assessed. Finally, participants were instructed that they would have a final test to assess their learning. Next, they read the science text and answered the multiple-choice questions one-by-one on Read&Learn at their own pace. Half of the students knew that they would not have access to the text once they clicked on the ‘Questions’ button (Unavailable text), whereas the other half knew they could refer back to the text to search for information (Available text). After answering each question, half of the students received feedback on the correctness of their response (KR feedback), whereas the other half received non-formative feedback messages (Control

feedback). Next, all the students had the option to access the EF message. Participants were never forced to access the EF messages, but they were encouraged to do so by informing them that this information may be helpful to understand the knowledge assessed. After viewing the feedback, the participants clicked a ‘Close feedback’ button and the question was presented with their answer choice selected. In the third session, participants solved a new set of short-answer questions in paper-and-pencil format. This final test covered the main knowledge assessed in the learning phase and was administered after a one-hour break to remove the short-term memory effects. The students had to answer the questions with the information they had learned during the previous session.

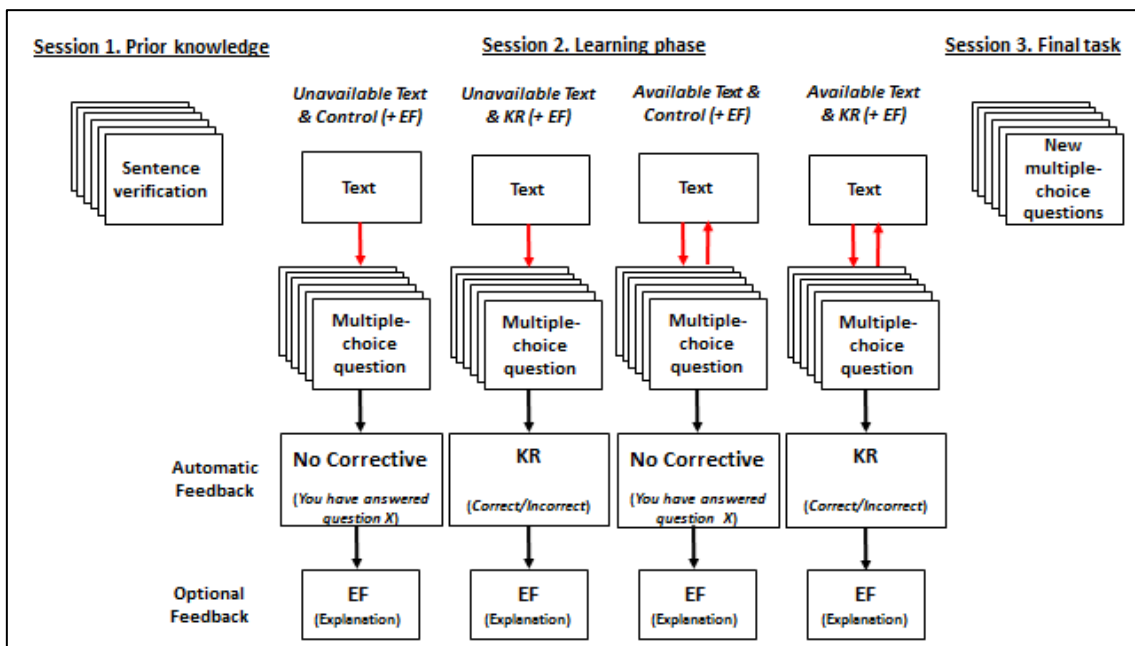


Figure 4. Schematic of the experimental procedure used for the Unavailable/Available text and the Control/KR feedback groups.

2.5. Online measures

For the learning assignment task, Read&Learn computed two main variables of interest: (1) *Performance on the assignment task* as the number of correct responses to the 12 multiple-choice questions, and (2) *EF access* as the number of times students decided to open the EF messages for the 12 questions. The latter measure was also computed conditional to the question-answering success (i.e., providing the correct or incorrect answer) in order to provide a deeper understanding of EF use.

3. RESULTS

3.1. Effects of KR feedback and text availability on accessing EF

The first aim was to determine to what extent KR feedback and text availability influenced students' decision to access additional EF during the learning assignment task. To address this aim we conducted a univariate ANCOVA with *EF access* as dependent variable, KR feedback (KR, control) and text availability (unavailable, available) as independent variables, and students' prior knowledge as covariate. Descriptive statistics are provided in Table 1. As predicted in Hypothesis 1, there was a main effect of text availability, $F(1, 134) = 4.56, p = .035, \text{partial } \eta^2 = .03$. Post-hoc analyses showed that students accessed EF messages more often in the unavailable-text ($M = 7.11, SD = 4.06$) than in the available-text ($M = 5.67, SD = 4.01$) condition, $t(134) = 2.14, p = .035$. KR feedback, however, did not affect the students' overall decision to access EF, $F(1, 134) = 2.32, p = .13$, which is inconsistent with Hypothesis 2. Further, there was no interaction between text availability and KR feedback, $F(1, 134) = 1.66, p = .20$, which does not support our Hypothesis 3 suggesting that having the text unavailable and receiving non-corrective feedback would trigger EF accesses more often. Also, the covariate, student's prior knowledge, did not affect the decisions to access EF, $F(1, 134) = 3.17, p = .08$.

Table 1. Means and Standard Deviations of the EF Access as a Function of Text Availability (Unavailable, Available) and KR Feedback (KR, Control)

	Unavailable		Available	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
KR	7.00	4.05	4.71	3.61
Control	7.24	4.13	6.65	4.21

Note. Feedback access max = 12

To fully examine our Hypothesis 2 (i.e., students who receive KR feedback would use EF mainly after failure), we explored whether *EF access* depended on the student's success to answer the questions by means of a mixed ANCOVA with KR feedback (KR, control) and text availability (unavailable, available) as between-subjects variables, question-answering success (correct, incorrect) as a within-subjects variable, and prior knowledge as covariate. Descriptive statistics are provided in Table 2. Results

showed a marginally-significant effect of question-answering success, $F(1, 134) = 3.74$, $p = .055$, $partial \eta^2 = .03$. Pairwise comparisons revealed that participants decided to access EF messages more often after providing an incorrect ($M = 3.47$, $SD = 2.32$) than a correct answer ($M = 2.93$, $SD = 2.67$), $t(134) = 2.23$, $p = .027$. As predicted in Hypothesis 2, there was a significant interaction for question-answering success and KR feedback (Figure 5), $F(1, 134) = 16.58$, $p < .001$, $partial \eta^2 = .11$. Simple effects analyses revealed that the increase in accessing the EF messages after providing incorrect answers was significant for the KR feedback group, $F(1, 70) = 20.69$, $p < .001$, $partial \eta^2 = .23$, but not for the control feedback group, $F(1, 67) = 1.48$, $p = .228$, $partial \eta^2 = .02$. No other effects were significant (all $ps > .05$).

Table 2. Means and Standard Deviations of the EF Access, as a Function of Question-Answering Success (Correct, Incorrect), Text Availability (Unavailable, Available), and KR Feedback (KR, Control)

	Correct Answers				Incorrect Answers			
	Unavailable		Available		Unavailable		Available	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
KR	2.86	2.65	1.54	2.17	4.14	2.19	3.17	2.33
Control	3.91	2.68	3.44	2.64	3.32	2.43	3.21	2.28

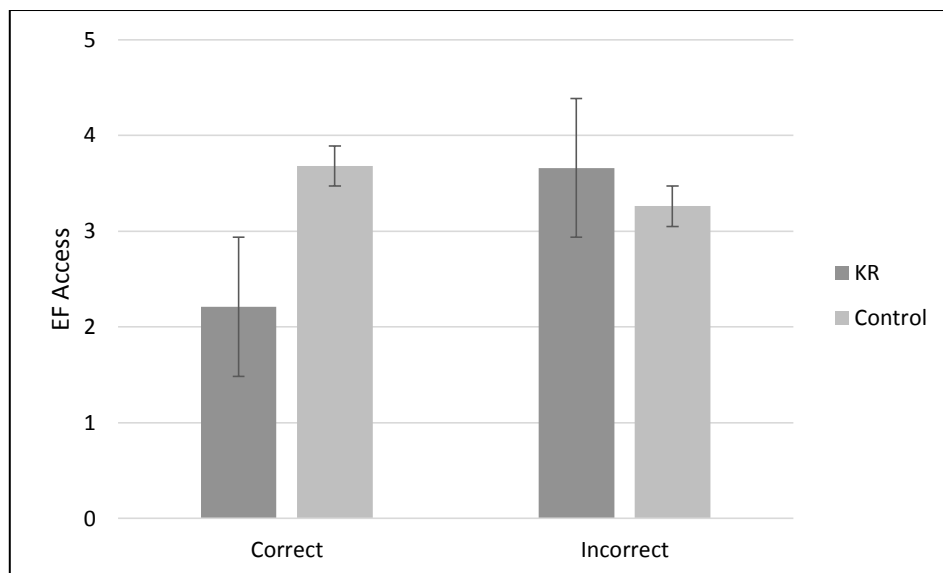


Figure 5. Number of decisions to access the EF messages after providing correct or incorrect responses for the KR and Control feedback groups

3.2. Effects of KR feedback and text availability on student's performance

The second aim was to determine to what extent KR feedback and text availability influenced the students' learning outcomes. To address this aim we conducted two separate univariate ANCOVAs with KR feedback (KR, control) and text availability (unavailable, available) as independent variables, prior knowledge as covariate, and performance outcomes (i.e., learning assignment task and the final test performance) as dependent variables. Descriptive statistics for the learning assignment task and the final test are provided in Table 3 and Table 4 respectively. For the *learning assignment task*, there was no effect of KR feedback, $F(1, 134) = 0.20, p = .657$, text availability, $F(1, 134) = 0.002, p = .963$, or interaction, $F(1, 134) = 0.89, p = .346$. Student's prior knowledge, however, was significantly related to their performance on the learning task, $F(1, 134) = 16.24, p < .001, partial \eta^2 = .11$. The same pattern of results was found for the *final test* performance. Results showed that there was no effect of KR feedback, $F(1, 134) = 0.001, p = .971$, text availability, $F(1, 134) = 0.31, p = .579$, or interaction, $F(1, 134) = 0.05, p = .831$. Student's prior knowledge, however, was significantly related to their performance on the final task, $F(1, 134) = 20.03, p < .001, partial \eta^2 = .13$. Although it is true that KR feedback and text availability do not influence the student's performance after a delay (i.e., final test), our findings do not provide support to our Hypothesis 4 suggesting that having the text available and receiving KR feedback would enhance student's performance during the learning task.

Table 3. Means and Standard Deviations of the Learning Assignment Task Score as a Function of Text Availability (Unavailable, Available), and KR Feedback (KR, Control)

	Unavailable		Available	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
KR	6.08	1.86	6.37	2.07
Control	6.44	2.00	6.18	1.87

Note. Learning assignment task score max = 12

Table 4. Means and Standard Deviations of the Final Test Score as a Function of Text Availability (Unavailable, Available), and KR Feedback (KR, Control)

	Unavailable		Available	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
KR	4.00	2.14	4.20	1.69
Control	3.97	1.51	4.09	1.91

Note. Final test score max = 8

3.3. Does EF access moderate the effect of prior knowledge on student's performance?

The current set of findings raised an additional question that we explored with moderation analyses. Recall that in the previous analyses, prior knowledge did not have a direct effect on EF access, probably due to controlling prior knowledge across conditions or to potential moderating effects of EF access. Thus, we explored whether *EF access* might influence the relation between students' prior knowledge and performance outcomes (i.e., learning assignment task and final test). We conducted two moderation analyses that included EF access as a moderator of the relation between student's prior knowledge and performance outcomes.

For the *learning assignment task*, the first model accounted for 15% of the variance in performance. Moderation was revealed by a significant interaction between prior knowledge and EF access ($\beta = -.02$, $t(135) = -2.06$, $p = .041$). As predicted in the second explanation of our Hypothesis 5, the relation between students' prior knowledge and performance is moderated by the number of decisions to access EF messages. Additionally, prior knowledge significantly predicted performance ($\beta = .15$, $t(135) = 4.10$, $p < .001$). However, EF access did not significantly predict performance by itself ($\beta = .05$, $t(135) = 1.31$, $p = .19$). To interpret the moderation effect we examined the simple slopes (Figure 6), which indicate that when participants accessed few times the EF messages, there was a significant positive relation between prior knowledge and learning assignment task performance ($\beta = .23$, $t(135) = 4.52$, $p < .001$). At the average value of EF access, there was also a significant positive relation ($\beta = .15$, $t(135) = 4.10$, $p < .001$). However, when participants accessed many times the EF messages, there was a non-significant relation between prior knowledge and learning assignment task performance ($\beta = .07$, $t(135) = 1.32$, $p = .19$).

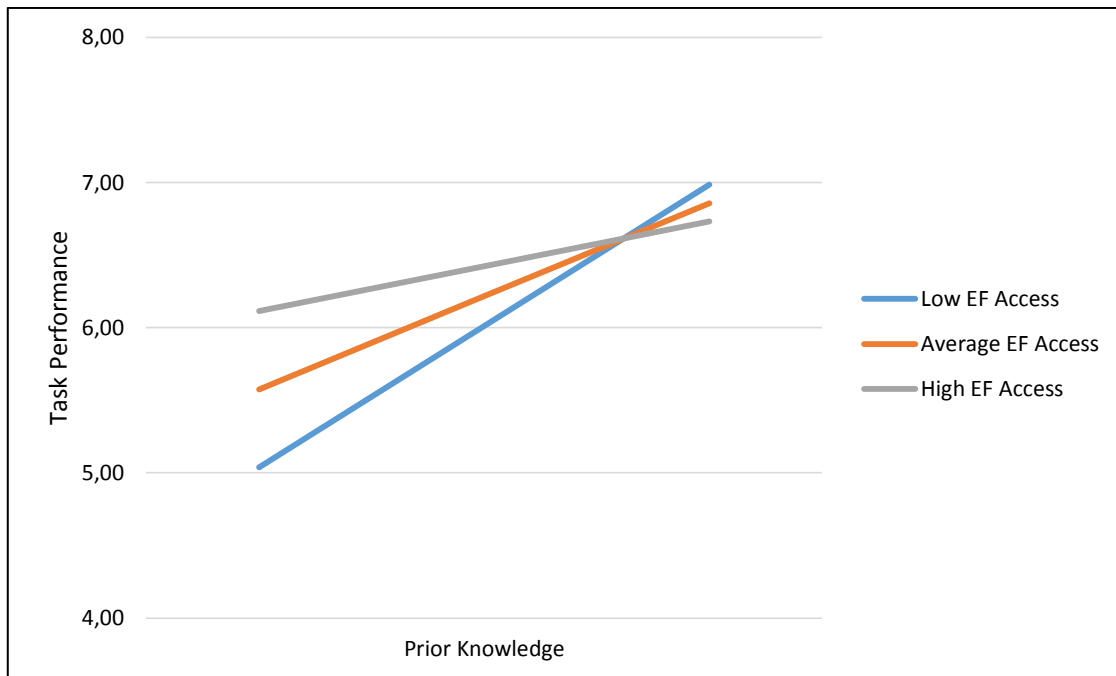


Figure 6. Moderation effect of EF access between students' prior knowledge and learning assignment task performance

For the *final test*, the second model accounted for 20% of the variance in performance. Moderation analysis revealed a marginally significant interaction effect between prior knowledge and EF access ($\beta = -.02$, $t(135) = -1.84$, $p = .069$). However, both prior knowledge ($\beta = .16$, $t(135) = 4.88$, $p < .001$) and EF access ($\beta = .09$, $t(135) = 2.58$, $p = .011$) significantly predicted final test performance. As predicted in the first explanation of our Hypothesis 5, EF and prior knowledge played independent roles when predicting the final test performance, although the interaction was marginally significant and simple slopes analyses showed the same trend as the slopes for the learning assignment task performance. Specifically, when participants accessed few times the EF messages, there was a positive relation between prior knowledge and performance. However, when participants accessed many times the EF messages, there was no relation between prior knowledge and performance.

4. DISCUSSION

Formative feedback is one of the core components in the design of digital learning environments (Shute & Rahimi, 2017; van der Kleij et al., 2015). Although feedback effectiveness has been widely investigated (e.g., Hattie & Gan, 2011; Shute,

2008), there is little research on how learners use feedback (Fox et al., 2014; Máñez et al., 2019; Timmers & Veldkamp, 2011) and how external factors influence the feedback use in digital environments (Lefevre & Cox, 2016). In this study we aimed to determine the extent to which two external sources of information (i.e., KR feedback and text availability) influence Secondary-school student's decision to use EF made of explanations when learning conceptual knowledge in the domain of physics. Given the recent expansion of computer-based learning environments (e.g., Graesser, Conley, & Olney, 2012; VanLehn, 2011), examining how external factors play a part in using additional EF is of practical importance.

Regarding the first aim, to determine the extent to which KR feedback and text availability influence the student's decision to use additional EF, results show that both external sources of information play specific roles in the way students engage in feedback use. Findings suggest that KR feedback has a mere informative value about the correctness of a student's response. Interestingly, the presence or absence of KR feedback did not influence the student's overall decisions to access EF. What influenced EF access was the presence of KR feedback after failure. In this case, students accessed EF more frequently after providing incorrect responses. Nevertheless, students who did not receive KR feedback (i.e., non-corrective feedback) had to rely on their metacognitive judgment about question-answering success which, as we know from the metacomprehension literature, is quite inaccurate (Bjork, Dunlosky, & Kornell, 2013; Glenberg & Epstein, 1985). In this case, students accessed EF equally often for both correct and incorrect responses. Regarding the text availability, results show that students decided to use EF more often when the text was unavailable, suggesting that the lack of text information while answering questions enhances the student's willingness to use additional EF. However, contrary to our predictions, no interaction between KR feedback and text availability came out. Students who answered the questions under the unavailable-text condition and non-corrective feedback did not access EF more often compared to the other conditions in which at least one external source of information (either the KR feedback or the text) was available.

With regard to the second aim, to determine the extent to which KR feedback and text availability influence the student's learning outcomes, results show that there are no differences among the experimental groups, which is consistent with recent research in the area of science learning (Maier et al., 2016). KR feedback was ineffective to boost learning, which confirms the limited role played by this type of

feedback (Van der Kleij et al., 2015). Likewise, text availability did not influence the students' performance, probably because of the question format. Multiple-choice questions depend on recognition memory (Ozuru et al., 2007) and provide multiple retrieval cues, thus reducing reliance on text availability (Ferrer et al., 2017). Further, no interaction between KR feedback and text availability came out. Students who answered the questions under available-text condition and received KR feedback did not increase their performance during the learning task as expected. This lack of effects may be related to the complexity of the knowledge assessed in the domain of physics. The percentages of correct responses were relatively low, so that we must be cautious about feedback effectiveness in the context of the task administered.

An additional finding of the present study that is worth noting is the moderating role of EF access between prior knowledge and the students' learning task performance. Specifically, students with low prior knowledge who are more likely to use the additional EF messages benefit more from this feedback. This finding is in line with previous findings on feedback effectiveness in mathematics (e.g., Fyfe et al., 2012; Krause et al., 2009). The learners' decision to seek for EF reduces the positive effects of prior knowledge by drawing on another source of information that can influence their learning. Also interesting is the finding that both prior knowledge and EF use played a role (albeit different) in the final test. Students who accessed the EF messages often were more likely to use the newly-acquired knowledge in subsequent learning scenarios. These findings suggest that EF messages can promote transfer of learning since the final test administered required transfer of the newly-acquired knowledge to new scenarios (Butler et al., 2013; Van der Kleij et al., 2015). This makes good sense in the context of the feedback literature (Bangert-Drowns et al., 1991; Hattie & Gan, 2011; Hattie & Timperley, 2007).

The present set of findings has several theoretical implications for models of learning that include external feedback as a core component. One such model is the LP-FILE model (Learner Processing of Feedback in Intelligent Learning Environments) that outlines how learners process feedback by means of noticing, decoding, and making sense of its content (Timms, DeVelle, & Lay, 2016). First, and directly relevant to the LP-FILE model, this experiment provides theoretical information about how Secondary-school students monitor and self-regulate the use of EF when answering questions from a science text in a digital environment. This decision, however, can be influenced by external sources of information, such as the presence of KR feedback and

text availability, as shown in the present study. Thus, identifying the extent to which these and any additional factors influence EF access is critically important. Second, our findings suggest that it is necessary to distinguish among different types of feedback. Whereas KR feedback is simple and can be processed effortlessly, EF is complex and requires effort, which involves an explicit decision to process its content. Complementarily, the value of KR feedback is merely to inform the learner about failure or success, whereas the EF may contribute to deep learning by reducing the gap between the current and the desired level of performance (van der Kleij et al., 2015). Therefore, KR feedback can be considered the gateway to focusing attention on EF after failure. Third, EF effectiveness seems to be different depending on the learner's prior knowledge level. Thus, students with low prior knowledge who use EF quite often may obtain learning gains more easily.

These findings also have several practical implications for the design of digital learning environments (e.g., Intelligent Tutoring Systems, e-textbooks or Massive Online Open Courses) in which formative feedback can be delivered in a timely manner. First, providing students with KR feedback along with EF is essential because it verifies student's responses and triggers EF use for incorrectly-answered questions. Therefore, digital environments should provide both corrective and elaborative feedback information (Lefevre & Cox, 2016), especially when it comes to learning complex knowledge (Tärning, 2018). Second, having the text available when answering multiple-choice questions seems not to enhance students' performance, although it may affect the student's willingness to use EF. Finally, promoting the use of EF among the students with lower levels of prior knowledge may increase their learning gains.

This study has several limitations we need to acknowledge, a few already addressing in ongoing research. First, we designed a one-hour learning session, which represents a short learning phase to obtain durable results. Therefore, including more texts and questions, as well as other science topics in a longer intervention would improve the feedback effectiveness and the generalizability of the findings. Second, it would be interesting to include another corrective feedback condition on the correct response (i.e., KCR feedback) to explore whether students access EF with the aim of improving their understanding or merely knowing the correct response. Even though we explicitly avoided mentioning the correct answer, some students might have accessed EF for this purpose. Finally, motivational factors may have influenced the students'

decision to access EF. Collecting data about the students' achievement goals may provide a better understanding of their feedback-seeking behavior.

Open questions aside, the present study is among the first in the literature that has focused on how Secondary-school students use computer-based EF when learning conceptual knowledge of physics, a science area in which young learners usually experience difficulties. Findings suggest that students monitor the use of EF in self-regulated learning situations (Butler & Winne, 1995), and that feedback use may vary depending on external factors such as the presence of KR feedback and text availability. Results also revealed that learners with lower levels of prior knowledge may benefit from using EF given that they engage in processing its content often. Therefore, this study advances our understanding about feedback use and feedback effectiveness when learning conceptual knowledge in digital learning environments.

References

- Agarwal, P. K., Karpicke, J. D., Kang, S. H., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology, 22*(7), 861-876. doi:10.1002/acp.1391
- 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
- Barrio-Cantalejo, I. M., Simón-Lorda, P., Melguizo, M., Escalona, I., Marijuán, M. I., & Hernando, P. (2008). Validación de la Escala INFLESZ para evaluar la legibilidad de los textos dirigidos a pacientes. *Anales del Sistema Sanitario de Navarra, 31*(2), 135–152.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. doi:10.1146/annurev-psych-113011-143823
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology, 39*(3), 206-219. doi:10.1016/j.cedpsych.2014.05.003
- Butler, A. C., Godbole, N., & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology, 105*(2), 290-298. doi:10.1037/a0031026
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281. doi:10.3102/00346543065003245
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential

- mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687-700. doi:10.1037/a0019452
- Dunlosky, J., & Rawson, K. A. (2015). Do students use testing and feedback while learning? A focus on key concept definitions and learning to criterion. *Learning and Instruction*, *39*, 32-44. doi:10.1016/j.learninstruc.2015.05.003
- Ferrer, A., Vidal-Abarca, E., Serrano, M. A., & Gilabert, R. (2017). Impact of text availability and question format on reading comprehension processes. *Contemporary Educational Psychology*, *51*, 404-415. doi:10.1016/j.cedpsych.2017.10.002
- Fox, J.-P., Klein Entink, R., & Timmers, C. (2014). The joint multivariate modeling of multiple mixed response sources: Relating student performances with feedback behavior. *Multivariate Behavioral Research*, *49*, 54-66. doi:10.1080/00273171.2013.843441
- Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, *104*(4), 1094–1108. doi:10.1037/a0028389
- Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 702-718.
- Graesser, A.C., Conley, M., & Olney, A. (2012). Intelligent tutoring systems. In K.R. Harris, S. Graham, and T. Urdan (Eds.), *APA Educational psychology handbook: Vol. 3. Applications to learning and teaching* (pp. 451-473). Washington, DC: American Psychological Association.
- Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249 – 271). New York, NY: Routledge.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Higgs, K., Magliano, J. P., Vidal-Abarca, E., Martínez, T., & McNamara, D. S. (2017). Bridging skill and task-oriented reading. *Discourse Processes*, 54(1), 19-39. doi:10.1080/0163853X.2015.1100572
- Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, 57, 219–232. doi:10.1007/BF03395573
- Kendeou, P., & O'Brien, E. J. (2018). Theories of text processing: A view from the top-down. In M. Schober, D. N. Rapp, & M. A. Britt (Eds.). *Handbook of discourse processes (2nd edition)* (pp. 19-33). New York: Routledge Publishing.
- 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
- Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, 19(2), 158 –170. doi:10.1016/j.learninstruc.2008.03.003
- Lefevre, D., & Cox, B. (2016). Feedback in technology-based instruction: Learner preferences. *British Journal of Educational Technology*, 47(2), 248-256. doi:10.1111/bjet.12227
- Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85-98. doi:10.1016/j.compedu.2015.12.002

- Máñez, I. (2018). *¿Influye la retroalimentación correctiva en el uso de la retroalimentación elaborada en un entorno digital?* Manuscript submitted for publication.
- Máñez, I., Vidal-Abarca, E., Kendeou, P., & Martínez, T. (2019). How do students process complex formative feedback in question-answering tasks? A think-aloud study. *Metacognition and Learning*. doi:10.1007/s11409-019-09192-w
- Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, 60(1), 95-109. doi:10.1016/j.compedu.2012.07.011
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22(3), 247-288.
- Mory, E. H. (2004). Feedback research revisited. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Erlbaum.
- 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
- OECD. (2010). *PISA 2009 results: What students know and can do: Student performance in reading, mathematics and science*. Paris: OECD Publishing.
- Ozuru, Y., Best, R., Bell, C., Witherspoon, A., & McNamara, D. S. (2007). Influence of question format and text availability on the assessment of expository text comprehension. *Cognition and Instruction*, 25, 399–438. doi:10.1080/07370000701632371

- Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, 47(6), 1525-1538. doi:10.1037/a0024997
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. doi:10.3102/0034654307313795
- Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33(1), 1-19. doi:10.1111/jcal.12172
- Smits, M., Boon, J., Sluijsmans, D. M. A., & Van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: Effects on learning as a function of prior knowledge. *Interactive Learning Environments*, 16(2), 183–193. doi:10.1080/10494820701365952
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. London: Routledge.
- Tärning, B. (2018). Review of feedback in digital applications. Does the feedback they provide support learning? *Journal of Information Technology Education: Research*, 17, 247-283. doi:10.28945/4104
- 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
- Timms, M., DeVelle, S., & Lay, D. (2016). Towards a model of how learners process feedback: A deeper look at learning. *Australian Journal of Education*, 60(2), 128-145. doi:10.1177/0004944116652912

- Van der Kleij, F. M., Eggen, T. J. H. M., Timmers, C. F., & Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, 58, 263–272. doi:10.1016/j.compedu.2011.07.020
- 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
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221. doi:10.1080/00461520.2011.611369
- 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

CAPÍTULO 6

CONCLUSIONES GENERALES

Responder preguntas a partir de textos expositivos es una tarea que los profesores suelen asignar a sus alumnos muy frecuentemente para facilitar la comprensión del texto y promover el aprendizaje de sus contenidos (Ness, 2011; Sánchez y García, 2015; Sánchez et al., 2010). Según el reciente modelo RESOLV (Britt et al., 2018), la lectura dirigida a metas puede conceptualizarse como una actividad de solución de problemas en la que los lectores usan textos con el objetivo de realizar tareas con fines específicos como, por ejemplo, responder preguntas. Estas tareas requieren que los estudiantes apliquen no sólo procesos de comprensión (e.g., realizar inferencias entre varios segmentos del texto), sino también habilidades específicas en el manejo de textos (e.g., cómo buscar la información en el texto o qué información buscar en función de las demandas de la pregunta) (OECD, 2010; Rouet, 2006; Rouet y Britt, 2011; Rouet et al., 2017; Vidal-Abarca et al., 2010). Este tipo de tareas ha recibido atención a lo largo de estos últimos años debido a la necesidad de comprender cómo afrontan los estudiantes el manejo de textos para alcanzar objetivos específicos (e.g., Anmarkrud et al., 2013; Cerdán et al., 2011; Gil et al., 2015; Mañá et al., 2017; McCrudden y Schraw, 2007; Roelle y Berthold, 2017; Vidal-Abarca et al., 2010). Al responder cada pregunta, el estudiante genera un modelo de tarea más o menos preciso que guiará la forma de contestar cada pregunta. Si el estudiante decide que es necesario buscar información, entonces empleará el texto para localizar información relevante que le ayude a responder correctamente la pregunta (McCrudden et al., 2011; McCrudden y Schraw, 2007). Si no lo considera necesario, entonces intentará responder la pregunta con la representación mental accesible en el momento de responder. A lo largo de este proceso, el estudiante genera un modelo de respuesta que puede actualizar cuando procesa el texto o la pregunta. Cuando el texto no está disponible, sin embargo, el estudiante ha de responder las preguntas con la representación mental que haya construido durante la lectura del texto antes de completar la tarea. Finalmente, el estudiante valida su modelo de respuesta cuando considera que es correcto o no puede mejorarlo, generando así expectativas sobre el resultado de su respuesta. Frecuentemente la tarea de responder preguntas a partir de textos tiene como objetivo último aprender determinado conocimiento conceptual.

Al responder las preguntas, agentes externos, como profesores o sistemas informáticos, suelen proporcionar a los estudiantes retroalimentación sobre sus respuestas para mejorar la comprensión de los materiales, corregir errores, y mejorar su rendimiento. La retroalimentación formativa proporciona al aprendiz información sobre

su rendimiento actual con respecto a un estándar para modificar sus pensamientos o comportamientos y así mejorar su aprendizaje (e.g., Hattie y Gan, 2011; Hattie y Timperley, 2007; Narciss, 2008; Shute, 2008). Gracias al rápido desarrollo de las nuevas tecnologías y al incremento en el número de dispositivos tecnológicos en las aulas de los centros educativos, cada vez es más común ver a estudiantes completar este tipo de tareas en entornos digitales, en los cuáles la retroalimentación formativa cobra especial importancia (e.g., Azevedo y Bernard, 1995; Mason y Bruning, 2001; Mory, 2004; Shute y Rahimi, 2017). Mientras que los profesores suelen proporcionar una retroalimentación demorada, tras contestar cada pregunta, y más centrada en los aspectos correctivos (i.e., fallo/error, o respuesta correcta), los sistemas digitales son capaces de proporcionar retroalimentación inmediata, tras cada pregunta, así como proporcionar algún tipo de EF específico tras responder, ya que son capaces de registrar datos sobre la interacción del alumno con los materiales en tiempo real. Los sistemas de tutorización inteligente como TuinLEC (Vidal-Abarca et al., 2014) o AutoTutor (Graesser et al., 2004; Olney, Graesser y Person, 2010), así como los libros de texto electrónicos, permiten proporcionar retroalimentación a los estudiantes en un medio digital. Además, son capaces de recoger información sobre el uso que hacen los estudiantes de esa información (e.g., *¿buscan retroalimentación? ¿cuánto tiempo dedican a procesarla?*). Sin embargo, estos sistemas requieren de investigación específica que analice cómo emplean la retroalimentación formativa los estudiantes al objeto de proporcionar una experiencia de aprendizaje eficaz.

La efectividad de la retroalimentación ha sido estudiada en una amplia variedad de contextos y actividades de aprendizaje en entornos digitales, desde simples tareas de memorización de vocabulario o definiciones, a tareas más complejas de comprensión y aprendizaje de conocimientos conceptuales (e.g., Butler et al. 2013; Corbalan et al., 2010; Golke et al., 2015; Lee et al., 2009; Llorens et al., 2016; Maier et al., 2016; Moreno, 2004; Murphy, 2007). Las tareas de comprensión y aprendizaje a partir de textos requieren procesos cognitivos de orden superior (e.g., realizar inferencias o aplicar los conocimientos aprendidos a nuevas situaciones). En estas tareas, la retroalimentación juega un papel muy importante, aunque su efectividad es muy variable y no todos los tipos de retroalimentación son igualmente efectivos. El reciente meta-análisis de Van der Kleij et al. (2015) sobre la efectividad de la retroalimentación en entornos digitales pone de manifiesto que los mensajes EF son los más efectivos, los KCR tienen un efecto menor, y los mensajes KR no mejoran el aprendizaje. De hecho,

los autores coinciden en afirmar que la retroalimentación no debe limitarse a informar únicamente del acierto/error o incluir la respuesta correcta, sino que es más útil proporcionar EF, especialmente cuando la tarea requiere procesos cognitivos de alto nivel (e.g., Butler et al., 2013; Hattie y Timperley, 2007; Kluger y DeNisi, 1996; Mory, 2004; Narciss, 2004; Shute, 2008).

Para que la retroalimentación formativa mejore el aprendizaje, los modelos teóricos más recientes han superado las limitaciones de los modelos conductistas de principios de siglo XX, y reconocen el papel activo de los aprendices durante su procesamiento, especialmente cuando se trata de EF (e.g., Bangert-Drowns et al., 1991; Hattie y Timperley, 2007; Narciss, 2013; Timms et al., 2016). Así, los estudiantes deben participar activamente en el procesamiento de su contenido al objeto de que la información sea integrada en su representación mental por medio de actualizaciones en sus conocimientos. En otras palabras, los estudiantes han de monitorizar activamente su rendimiento mediante la búsqueda de correspondencia entre su nivel de comprensión y el nivel proporcionado en la retroalimentación para ajustar su representación al estándar proporcionado (e.g., Bangert-Drowns et al., 1991; Butler y Winne, 1995). Para ello necesitan implementar procesos cognitivos y metacognitivos (Carless y Boud, 2018), por lo que la retroalimentación puede tener un significado muy diferente en función de la interpretación que haga cada estudiante (Black y Wiliam, 2009). No obstante lo anterior, no todos los tipos de retroalimentación requieren el mismo nivel atencional. Mientras que los mensajes de retroalimentación correctiva KR y KCR requieren pocos recursos cognitivos para ser procesados, los mensajes EF requieren un procesamiento activo que permita a los estudiantes analizar sus errores a la luz de la información recibida. A pesar de la vasta investigación realizada sobre su efectividad, aún se desconocen los mecanismos relacionados con el procesamiento de la retroalimentación formativa en entornos digitales (Van der Kleij et al., 2012). En este sentido, y siguiendo los modelos recientes, consideramos que existen dos fases en el procesamiento de la EF que incluye algún tipo de retroalimentación correctiva. En la primera fase, los participantes tratan de verificar o refutar el modelo de respuesta generado durante la fase de resolución de la tarea. En la segunda fase, los participantes han de decidir si necesitan revisar su conocimiento, por lo que aplican procesos de comprensión, monitorización y autorregulación que pueden producir actualizaciones en el modelo de respuesta validado. Así, para que la EF dé lugar a cambios en los conocimientos aprendidos, los estudiantes han de estar dispuestos y ser capaces de comprender la

información en el contexto de tarea asignado (Timmers y Veldkamp, 2011; Timms et al., 2016).

La EF puede contener información muy diversa. En los dos primeros estudios presentados, el programa informático Read&Learn proporcionó a estudiantes de 3er curso de secundaria mensajes EF tras responder cada pregunta de comprensión y haber seleccionado la información textual relevante. Estos mensajes EF estaban formados por retroalimentación KR y KCR sobre la precisión de sus respuestas y KR sobre la precisión para localizar información relevante, así como pistas para mejorar las ejecuciones posteriores. Además, Read&Learn les permitió acceder voluntariamente a la pantalla del texto con retroalimentación KCR que les informaba sobre el estándar de información relevante que debían haber seleccionado. En los dos últimos estudios, sin embargo, Read&Learn proporcionó a los estudiantes mensajes EF que contenían explicaciones sobre los contenidos de un texto expositivo de física, tras recibir retroalimentación correctiva KR o KCR automáticamente. Aunque estos y otros tipos de EF pueden ser administrados automáticamente en entornos digitales de aprendizaje, esto no significa que los estudiantes decidan utilizarlos también automáticamente, sino que requiere de un proceso activo de toma de decisiones que puede estar influido por variables contextuales (e.g., presencia de retroalimentación correctiva o disponibilidad de los materiales) y características personales (e.g., competencia lectora o nivel de conocimiento previo). Por tanto, resulta esencial realizar investigación específica que analice la forma en que estudiantes jóvenes usan y procesan los mensajes EF en entornos digitales, objetivo central de este trabajo.

El objetivo general del presente trabajo consistía en examinar cómo interactúan los estudiantes de 3er grado de secundaria con unos mensajes de retroalimentación EF proporcionados por el sistema informático Read&Learn tras responder preguntas de comprensión y aprendizaje, así como sus efectos en el rendimiento, teniendo en cuenta la competencia lectora y el nivel de conocimiento previo de los estudiantes. Para ello realizamos cuatro estudios con objetivos específicos. Mientras que los dos primeros estudios nos sirvieron para examinar el uso y procesamiento de una EF específica en una tarea que requería comprender ideas conceptuales de unos textos expositivos no académicos, los dos últimos estudios sirvieron para examinar la decisión de los estudiantes para acceder voluntariamente a unos mensajes EF con explicaciones sobre los contenidos de un texto expositivo de ciencias. Estas decisiones fueron registradas por Read&Learn.

El objetivo del estudio 1 consistió en analizar el efecto de una EF específica que contenía información sobre el rendimiento al responder preguntas de comprensión y seleccionar información relevante (i.e., informaba sobre qué información seleccionada era relevante e irrelevante y, además, presentaba el estándar con la información relevante al revisar el texto antes de pasar a la siguiente pregunta). Para ello pedimos a los participantes que respondieran un conjunto de preguntas de comprensión de unos textos expositivos no académicos y que seleccionaran la información textual que consideraban relevante para responder cada pregunta. Además, analizamos en qué medida los estudiantes que recibieron EF atendieron a esos mensajes por encima de un grupo control sin retroalimentación, controlando la influencia de la competencia lectora. Los resultados sugieren que la EF puede mejorar el rendimiento de los estudiantes, y que estos están dispuestos a dedicar ligeramente algo más de tiempo a procesar la EF en comparación con el grupo control. Estos resultados nos llevaron a preguntarnos cómo se procesa la EF, es decir, qué operaciones mentales aplican los estudiantes cuando reciben estos mensajes de retroalimentación al responder preguntas de comprensión. Éste fue el objetivo del estudio 2. Los resultados de este segundo estudio sugieren que los estudiantes están principalmente interesados en verificar sus respuestas y conocer la respuesta correcta, que emplean la EF para realizar comparaciones poco elaboradas entre su modelo de respuesta y el estándar proporcionado, y que en contadas ocasiones tratan de construir significado a partir de esta EF. Además, no encontramos diferencias en el procesamiento de la retroalimentación entre los estudiantes con alta y baja competencia lectora. Este conjunto de resultados nos llevó a realizar una serie de modificaciones en el diseño de los dos últimos estudios.

El objetivo del estudio 3 consistió en conocer el efecto de la presencia de la retroalimentación correctiva KR y KCR en el acceso voluntario a mensajes EF que contenían explicaciones sobre los contenidos evaluados. En este caso empleamos un texto expositivo de ciencias y unas preguntas para fomentar el aprendizaje de sus contenidos, controlando el nivel de conocimiento previo de los estudiantes. Los resultados sugieren que la presencia de retroalimentación KCR limita el uso de la EF significativamente, aunque no encontramos diferencias en el rendimiento de los estudiantes. En el estudio 4 exploramos el efecto de la retroalimentación KR y la disponibilidad del texto en el acceso voluntario a esos mensajes EF, ya que éstas son dos variables del contexto de tarea que pueden jugar un papel determinante en el uso de la EF. Este estudio nos permitió corroborar cómo la retroalimentación KR permite a los

estudiantes verificar o refutar su modelo de respuesta y utilizar la EF tras responder incorrectamente, así como el incremento en el uso de la EF cuando el texto no es una fuente de información disponible durante la tarea. Por último, los resultados también sugieren que los estudiantes con bajo conocimiento previo que deciden utilizar la EF frecuentemente reducen el impacto de su conocimiento previo en la tarea.

El conjunto de resultados obtenidos en los cuatro estudios nos permite formular varias conclusiones generales relacionadas tanto con el uso y procesamiento de la EF como con su efectividad en las tareas de responder preguntas de comprensión y aprendizaje en entornos digitales.

Por lo que respecta al uso y procesamiento de la retroalimentación administrada, una de las primeras conclusiones que se derivan del conjunto de estudios presentados es que los estudiantes procesan los mensajes EF que contienen retroalimentación correctiva de una forma particular. Parece que los estudiantes emplean la retroalimentación para verificar o refutar su modelo de respuesta, tal y como postulamos en la primera fase del modelo sobre procesamiento de la retroalimentación. Cuando las respuestas de los estudiantes son erróneas y son informados de ello, los estudiantes emplean la EF para realizar modificaciones en el modelo de respuesta, tal y como postulamos en la segunda fase del modelo. Estas actualizaciones suelen ser simples y consisten en la codificación de la respuesta correcta cuando se proporciona retroalimentación KCR. Así, los estudiantes monitorizan continuamente el uso que hacen de la EF que contiene retroalimentación correctiva al objeto de reducir la discrepancia entre su nivel de comprensión y el estándar (e.g., Hattie y Timperley, 2007; Mason y Bruning, 2001). Para ello, los estudiantes aplican operaciones cognitivas (e.g., parafrasear o elaborar) y metacognitivas (e.g., comparar la ejecución con el estándar, evaluar la propia comprensión) específicas que producen modificaciones en los modelos de respuesta que verifican o refutan al recibir retroalimentación.

En segundo lugar, el segundo estudio de este trabajo examinó la prevalencia de las operaciones mentales que aplican los estudiantes al recibir la EF administrada en el primer estudio mediante el uso de la técnica de *pensamiento-en-voz-alta*. Los resultados mostraron que los estudiantes focalizan principalmente su atención en conocer la corrección de sus respuestas y la respuesta correcta, dedican menos recursos a conocer la precisión con la que evalúan la relevancia de la información textual, y aún menos a conocer el estándar con la información relevante a la que deberían atender. El análisis de las verbalizaciones mostró que los estudiantes emplean la retroalimentación para

evaluar su ejecución y su comprensión (i.e., procesos de monitorización), para construir significado en ocasiones (i.e., procesos cognitivos de comprensión), y para regular el uso de la información proporcionada en los mensajes (i.e., procesos de autorregulación). Entre los procesos de monitorización, cabe destacar la elevada frecuencia de las comparaciones entre la propia ejecución y el estándar proporcionado. La inclusión de la respuesta correcta y la información relevante pudo ser clave para fomentar este proceso de monitorización en el que el estudiante intenta comprender la relación entre el rendimiento actual y el deseado. Asimismo, los estudiantes respondieron afectivamente, aplicaron procesos atribucionales y evaluaron su propia comprensión y memoria. Entre los procesos cognitivos de comprensión, destacan las paráfrasis y la lectura de información, superando al número de elaboraciones realizadas, por lo que podemos concluir que los estudiantes construyen significado de forma superficial a partir de la información contenida en este tipo de mensajes EF. Además, observamos cómo monitorizan los estudiantes el uso de la retroalimentación. Debido a que la EF no se procesa automáticamente y los estudiantes pueden omitir parte de la información proporcionada, parece que ofrecer a los estudiantes la oportunidad de acceder opcionalmente a mensajes EF es un elemento a considerar en el diseño de entornos digitales de aprendizaje (e.g., Llorens et al., 2016; Timmers y Veldkamp, 2011).

Los estudios de esta tesis han evidenciado que usar la EF consiste en una decisión metacognitiva que varía en función de factores externos como la presencia de retroalimentación correctiva y la disponibilidad del texto, dos variables a considerar en el diseño de entornos digitales de aprendizaje. En el caso de la disponibilidad de los textos, los resultados sugieren que la ausencia del texto durante la tarea aumenta la disposición de los estudiantes a usar la EF, probablemente debido a que se trata de la única información accesible con la que comprender los contenidos evaluados. En el caso de la retroalimentación, la EF suele proporcionarse junto con retroalimentación correctiva KR o KCR (Van der Kleij et al., 2015). Mientras que la retroalimentación KR es simple y puede ser procesada sin prácticamente esfuerzo, la retroalimentación KCR y EF requieren esfuerzo consciente, lo que implica una decisión explícita de procesar su contenido. Como sugiere la investigación previa, la retroalimentación correctiva permite monitorizar el uso de la EF, ya que sirve para verificar o refutar el modelo de respuesta y dirigir los recursos atencionales del estudiante hacia la EF cuando se responde incorrectamente (e.g., Fox et al., 2014; Timms et al., 2016). Sin embargo, la ausencia de retroalimentación correctiva obliga a los estudiantes a confiar en su criterio

metacognitivo sobre la precisión de sus respuestas para buscar EF adicional. Como sugieren los resultados, los estudiantes que no reciben retroalimentación correctiva tienden a emplear la EF de forma similar después de proporcionar respuestas correctas e incorrectas, lo que sugiere que son imprecisos a la hora de monitorizar la precisión de sus respuestas (Bjork, Dunlosky y Kornell, 2013; Glenberg y Epstein, 1985). En este caso, parece que los estudiantes emplean la EF para intentar encontrar pistas que les permitan identificar la respuesta correcta.

En el caso de la retroalimentación correctiva KR y KCR, parece que esta información puede interferir en el uso de la EF, seguramente porque muchos estudiantes están interesados en conocer la corrección de sus respuestas, más que en alcanzar un aprendizaje profundo. Los resultados del tercer y cuarto estudio confirman que la retroalimentación KR permite contrastar las expectativas de rendimiento de los estudiantes, juzgar su nivel de comprensión y ser conscientes de los conocimientos erróneos previamente aprendidos, siendo la puerta de acceso al uso de la EF tras responder incorrectamente. La retroalimentación KCR, además de verificar las expectativas de rendimiento de los estudiantes, les permite modificar sus conocimientos (parcialmente) erróneos por medio de la codificación del modelo de respuesta correcto. Parece que los estudiantes consideran innecesario invertir esfuerzos en procesar la EF una vez recibida retroalimentación KCR. De hecho, la decisión de no consultar la EF aumenta significativamente cuando se conoce la respuesta correcta. Es probable que la presencia de retroalimentación KCR en los dos primeros estudios redujera los procesos cognitivos de construcción de significado. Así, los resultados sugieren que verificar una respuesta permite a los estudiantes confirmar su conocimiento y corregir posibles imprecisiones metacognitivas (Butler et al., 2008), mientras que refutar una respuesta parece desafiar su cognición y les lleva a dedicar más recursos para comprender el error (Maier et al., 2016). Este conjunto de resultados sugiere que el principal objetivo de los estudiantes al resolver este tipo de tareas consiste principalmente en proporcionar la respuesta correcta (Farr et al., 1990; Rouet et al., 2017; Rupp et al., 2006; Vidal-Abarca et al., 2010).

Estos resultados apoyan los supuestos planteados en el modelo RESOLV para situaciones de lectura (Rouet et al., 2017) y en el modelo KReC para la revisión de conocimientos incorrectos o *misconceptions* (Kendeou y O'Brien, 2014). Consideramos que estos supuestos podrían ajustarse al uso y procesamiento de la retroalimentación en el contexto de tarea planteado, consistente en responder preguntas de comprensión y

aprendizaje a partir de textos. Así, es necesario que los estudiantes a) tomen la decisión de procesar la retroalimentación (i.e., procesamiento activo), b) adapten su comportamiento en función de la información recibida (i.e., comportamiento adaptativo), c) realicen una evaluación del coste cognitivo y emocional en relación con los beneficios que podrían obtener (i.e., recursos limitados), d) realicen juicios sobre la percepción de haber proporcionado la respuesta correcta (i.e., una cuestión de certeza en la respuesta), e) se alcance un umbral que desencadene unas conductas u otras en el procesamiento de la retroalimentación (i.e., alcanzar el umbral), y f) comprendan la información de los mensajes en el contexto de tarea (i.e., interpretar la retroalimentación). Además de estos supuestos, consideramos que los resultados obtenidos también proporcionan apoyo a los supuestos formulados por Kendeou y O'Brien (2014) sobre la revisión del conocimiento erróneo o *misconceptions*. Así, los resultados muestran que la información previamente aprendida por el estudiante puede interferir con el aprendizaje de la información de la retroalimentación (i.e., principio de codificación). Esa información aprendida puede (re)activarse en la memoria de trabajo automáticamente (i.e., principio de activación pasiva). Si la información incorrecta y la información correcta están activas en la memoria de trabajo al mismo tiempo se puede producir la revisión de conocimientos (i.e., principio de co-activación). Una vez que la información correcta de la retroalimentación entra en contacto con el conocimiento erróneo, ésta puede integrarse con ese conocimiento impreciso en la memoria de trabajo (i.e., principio de integración). Una vez procesada la retroalimentación, esta información compite con el modelo de respuesta creado por el estudiante; si la información correcta de la retroalimentación incrementa su activación y, al mismo tiempo, disminuye la activación del modelo de respuesta erróneo, la nueva información pasa a formar parte de la representación mental de los conocimientos aprendidos (i.e., principio de competición de la activación). Cuando se dan estos supuestos, asumimos que el modelo de respuesta activo en la memoria de trabajo de los estudiantes durante el procesamiento de la retroalimentación puede ser actualizado, y así producir cambios en su representación mental de los conocimientos evaluados.

Respecto a las diferencias individuales, los estudios presentados muestran resultados diversos. Mientras que los dos primeros estudios no encuentran diferencias significativas en los procesos cognitivos y metacognitivos aplicados durante el procesamiento de la EF, el cuarto estudio sugiere que las diferencias en la efectividad de la EF pueden variar según el uso que realicen estudiantes con diferente nivel de

conocimiento previo. En general, los estudiantes con alta o baja competencia lectora, medida como la habilidad general para responder preguntas de comprensión de un texto disponible, tienden a centrarse en la corrección de sus respuestas de forma similar. Aunque los resultados del cuarto estudio sugieren que el uso de la EF tiene una influencia positiva general en el rendimiento de los estudiantes en una tarea administrada tras un periodo de demora, parece que su eficacia en la propia tarea de aprendizaje varía según el nivel de conocimiento previo del alumnado. Así, los estudiantes con bajos niveles de conocimiento previo en el área de ciencias que están dispuestos a utilizar la EF frecuentemente parecen beneficiarse más de esta retroalimentación, mientras que los estudiantes con altos niveles de conocimiento previo se benefician menos.

Relacionado con esta última idea sobre la efectividad de la retroalimentación, encontramos resultados dispares en el conjunto de estudios presentados. Los resultados del primer estudio sugieren que la EF que incluye información sobre la precisión de los estudiantes tanto para responder preguntas de comprensión como para seleccionar información relevante, parece mejorar la eficiencia de los estudiantes en este tipo de tareas que requieren usar documentos para responder preguntas (e.g., Llorens et al., 2015; Llorens et al., 2016). Parece que proporcionar información específica sobre la ejecución del estudiante incrementa la eficacia de la retroalimentación en tareas que implican procesos cognitivos de orden superior (i.e., Hattie y Timperley, 2007; Kluger y DeNisi, 1996; Mason y Bruning, 2001; Mory, 2004; Narciss, 2004; Shute, 2008; van der Kleij et al., 2015). Así, esta EF puede incrementar el rendimiento de los estudiantes al responder un mayor número de preguntas correctamente y mejorar la precisión con la que descartan información no relevante, dedicando aproximadamente el mismo tiempo a la tarea y la búsqueda de información que aquellos estudiantes que no reciben retroalimentación formativa. Así, la EF puede influir en la comprensión del texto durante la propia tarea, ya que los mensajes proporcionan información específica para superar los problemas de comprensión. A pesar de estos beneficios en la selección de información irrelevante, parece que la EF administrada es insuficiente para mejorar los procesos de monitorización relacionados con la selección de información relevante. En los dos últimos estudios, sin embargo, no encontramos diferencias entre los distintos tipos de retroalimentación administrados, lo que es consistente con algunas de las investigaciones previas en el aprendizaje de conocimientos de ciencias (Maier et al., 2016). Esta ausencia de efectos pudo estar relacionada con el uso de la EF. Así, aquellos

estudiantes que reciben retroalimentación KR tienden a emplear la EF en igual o menor medida que aquellos que no reciben retroalimentación correctiva. Sin embargo, aquellos que reciben retroalimentación KCR tienden a emplear la EF en menor medida. Por tanto, puede que las diferencias en el uso de la EF en función de la retroalimentación correctiva recibida expliquen la ausencia de efectos en el rendimiento final. La retroalimentación KCR y EF pudo haber contribuido al aprendizaje de los estudiantes, ya que esa información puede reducir la distancia entre el nivel de rendimiento actual y el deseado (Van der Kleij et al., 2015). En definitiva, los cuatro estudios que componen esta tesis han permitido avanzar en la comprensión acerca del uso y procesamiento de la retroalimentación formativa que hacen los estudiantes de secundaria en un entorno digital.

Limitaciones

A pesar de las contribuciones teóricas y prácticas que se derivan del conjunto de estudios presentados, este trabajo no está exento de limitaciones que abren el camino a futuras líneas de investigación. En primer lugar, las tareas administradas no estuvieron integradas dentro de una asignatura del currículum escolar, por lo que la ejecución de los estudiantes no tuvo una repercusión directa en su rendimiento académico. En este sentido, las tareas conllevan generalmente un riesgo bajo (i.e., *low stakes tasks*), por lo que los estudiantes suelen dedicar menos esfuerzos (e.g., Kupiainen, Vainikainen, Marjanen y Hautamäki, 2014), y el procesamiento de la retroalimentación podría verse afectado (Golke et al., 2015). Si los estudiantes consideran que las tareas administradas no tienen una repercusión directa en su aprendizaje, su nivel de implicación puede ser bajo, y la retroalimentación podría emplearse únicamente para verificar o refutar sus respuestas y conocer la respuesta correcta. Una posible vía para fomentar el uso de la EF en este tipo de tareas de bajo riesgo consistiría en contextualizar las tareas para asegurarse de que la EF se recibe como una información útil que merece ser procesada.

En segundo lugar, la comprensión de los materiales administrados podría requerir un periodo más largo de aprendizaje (Bransford, 2000). En nuestro caso, la fase de aprendizaje fue corta, y el periodo de demora entre tareas relativamente breve. Si el objetivo es obtener beneficios a largo plazo, es posible que un entrenamiento con retroalimentación formativa requiera de un mayor tiempo de exposición a los materiales, por lo que habría que examinar en qué medida es necesario incrementar la cantidad y la duración de la fase de aprendizaje.

En tercer lugar, las preguntas fueron administradas en formato de elección múltiple. Aunque este formato facilita la corrección de las respuestas a los sistemas informáticos como Read&Learn, no tiene en cuenta los procesos de construcción de respuestas propios de las preguntas en formato abierto. Desarrollos recientes en el área de los tutores inteligentes de aprendizaje permiten la corrección de preguntas abiertas, manteniendo un discurso con el propio estudiante durante la resolución de las tareas (e.g., Graesser et al., 2004) y proporcionando retroalimentación ajustada a la precisión de sus respuestas (Graesser, 2016; Nye et al., 2014), lo que resulta imposible para la versión actual de Read&Learn, la cual no permite realizar un análisis semántico de respuestas abiertas y proporcionar la retroalimentación formativa correspondiente.

En cuarto lugar, los estudiantes que participaron en los estudios no tuvieron la oportunidad de emplear la EF en un segundo intento al responder cada pregunta. Aunque proporcionar una segunda oportunidad después de responder incorrectamente podría haber incrementado la sensación de utilidad de la EF, investigaciones recientes en el área de comprensión lectora sugieren que los estudiantes hacen un uso limitado de los mensajes EF que se administran tras responder incorrectamente, incluso cuando disponen de un segundo intento para responder (Golke et al., 2015). Puede que los estudiantes utilicen ese intento para validar la segunda alternativa considerada más plausible, en lugar de emplear la EF para corregir errores y actualizar su modelo de respuesta. Asimismo, los dos primeros estudios no incluyeron una situación de transferencia sin retroalimentación formativa, por lo que los estudiantes tal vez valoraran la EF como poco útil. No obstante, no se disponen de datos sobre la utilidad percibida de la EF recibida.

En quinto lugar, nuestros estudios han examinado el uso de la EF únicamente en dos dominios de conocimiento. Mientras que los dos primeros estudios contemplan el uso y procesamiento de la EF en una tarea de competencia lectora propia del área de lengua, los dos siguientes examinan la decisión de usar la EF en el dominio de física en función de la retroalimentación correctiva y la disponibilidad del texto. Además, los mensajes EF empleados en los dos primeros estudios son diferentes de los mensajes EF empleados en los dos últimos. Por tanto, cabe plantearse si los resultados encontrados en ambas áreas son generalizables a otros dominios de conocimientos como, por ejemplo, la comprensión y aprendizaje de conocimientos en las áreas de historia o geografía. Además, los dos últimos estudios han analizado la decisión de utilizar los mensajes EF voluntariamente, pero no han explorado las operaciones mentales que

aplican los estudiantes. Igualmente, cabe señalar que los dos últimos estudios no contaron con una condición control pura sin ningún tipo de retroalimentación.

Finalmente, los estudios presentados en esta tesis se han centrado en estudiantes de 3er curso de secundaria, por lo que sería conveniente explorar el uso y procesamiento de la retroalimentación formativa por estudiantes de otros cursos académicos y de otras etapas educativas. Si nos centramos en las variables individuales exploradas, hemos de señalar que la competencia lectora, tal y como fue medida en los dos primeros estudios, no influyó decisivamente en el uso y procesamiento de la EF proporcionada. En ambos casos se midió la habilidad general de los estudiantes para responder preguntas de comprensión de un texto disponible. Como se explicó en la introducción, esta habilidad es crucial para formar el modelo de tarea y buscar información relevante para resolver las preguntas (e.g., Gil et al., 2015; Vidal-Abarca et al., 2010). Sin embargo, parece no estar directamente relacionada con el procesamiento de los mensajes EF administrados. Es posible que aspectos cognitivos más básicos de la lectura, como la velocidad de codificación o el nivel de vocabulario, expliquen el procesamiento de la retroalimentación de forma más precisa. Además, los estudios presentados en esta tesis tampoco contemplan otros factores intrapersonales que podrían afectar al uso de la EF como, por ejemplo, aspectos metacognitivos como la habilidad de los estudiantes para refutar conocimientos erróneos, o aspectos motivacionales como la motivación intrínseca o las metas de logro (e.g., Kluger y DeNisi, 1996; Narciss, 2013; Nicol y Macfarlane-Dick, 2006; Shute, 2008).

Futuras direcciones

Este trabajo abre nuevas líneas de investigación relacionadas con el uso y procesamiento de la retroalimentación formativa en entornos digitales, así como su repercusión en el rendimiento de los estudiantes cuando responden preguntas de comprensión y aprendizaje. En primer lugar, la comprensión y aprendizaje de conocimientos conceptuales complejos requiere una práctica extensa y una contextualización de las tareas asignadas. Por tanto, una línea de investigación futura podría dirigirse, por una parte, a incrementar la duración del entrenamiento, y por otra, a integrar las tareas asignadas en las materias curriculares. Aunque los resultados del primer estudio nos han permitido comprobar que la EF es capaz de producir mejoras incluso durante la propia realización de la tarea, cabe plantearse cuál es el periodo de

entrenamiento necesario para que la EF produzca modificaciones significativas en las representaciones mentales de los estudiantes a largo plazo.

En segundo lugar, sería útil conocer el papel de la EF en función del formato de las preguntas y el tipo de tareas administradas. Como se ha señalado anteriormente, las preguntas fueron administradas en formato de elección múltiple, por lo que los estudiantes no tuvieron que construir ninguna respuesta. Además de evaluar la efectividad de la EF al responder preguntas en formato abierto, también sería conveniente explorar si la EF facilita la comprensión de los estudiantes en otro tipo de tareas (e.g., preguntas que requieran completar información). Recientes desarrollos a nivel computacional permiten administrar este tipo de tareas, así como proporcionar retroalimentación en función de varios parámetros relacionados tanto con el rendimiento como con las conductas estratégicas de los estudiantes.

Otras líneas de investigación futura podrían explorar el uso de la EF en función de diferentes variables del contexto de tarea. Habitualmente, los sistemas de tutorización inteligentes incluyen avatares que dialogan con los participantes, pidiéndoles que resuelvan tareas y proporcionándoles retroalimentación (Graesser et al., 2004; Olney et al., 2010; Vidal-Abarca et al., 2014). Así, sería interesante examinar el papel que juegan los agentes virtuales al proporcionar retroalimentación formativa en los entornos digitales (e.g., D'Mello, Craig, Fike y Graesser, 2009; Jonhson y Lester, 2016; Johnson, Rickel y Lester, 2000). También convendría analizar otras variables del contexto de tarea, como el momento en que son proporcionados los mensajes de retroalimentación (i.e., *timing*), las restricciones de tiempo, o la presencia de un segundo intento después de responder incorrectamente. Respecto al *timing*, los sistemas informáticos permiten proporcionar mensajes EF tras responder cada pregunta, después de responder un conjunto determinado de preguntas, o al finalizar el test. Las restricciones de tiempo, por su parte, pueden afectar a la toma de decisiones de los estudiantes (Thiede y Dunlosky, 1999), por lo que restringir el tiempo de la tarea puede influir en la disposición de los estudiantes para procesar la EF. Igualmente, es posible que los estudiantes estén más interesados en usar la EF si disponen de un segundo intento para responder. Este tipo de manipulaciones son relativamente sencillas en entornos digitales y pueden ofrecer información útil para diseñar tareas de forma más efectiva.

En tercer lugar, se pueden formular algunas líneas de investigación sobre el uso y la efectividad de la retroalimentación formativa al responder preguntas de

comprensión y aprendizaje en entornos digitales. Nuevos estudios son necesarios para examinar cómo diferentes tipos de EF ayudan a refutar los modelos de respuesta contruidos durante el proceso de responder las preguntas. En los dos últimos estudios empleamos unos mensajes EF que permitían al alumno hacer inferencias sobre la respuesta correcta, pero sin expresarla explícitamente. Sería conveniente analizar si mensajes EF con explicaciones sobre el error facilitan las actualizaciones en el modelo de respuesta de los estudiantes más fácilmente, modificando así sus representaciones mentales de los conocimientos aprendidos. Igualmente, sería interesante estudiar hasta qué punto la retroalimentación KCR es capaz de beneficiar el aprendizaje de los estudiantes en estas tareas.

Una de las líneas de investigación que mayor repercusión puede tener en los próximos años hace referencia a la enseñanza de habilidades específicas para utilizar la retroalimentación formativa eficazmente. Los estudiantes necesitan ser competentes para extraer el significado de estos mensajes, por lo que necesitan reconocer y valorar positivamente la retroalimentación recibida, realizar juicios sobre la calidad de sus respuestas, controlar sus emociones, y tomar decisiones activamente sobre el procesamiento de la información (Carless y Boud, 2018). Por tanto, parece que los estudiantes tal vez necesiten una instrucción específica para aprender a utilizar la retroalimentación formativa en el contexto de tarea asignado, ya que para beneficiarse de esta información no basta con estar dispuesto a procesarla, sino que, además, es necesario disponer de habilidades para comprender su contenido correctamente en el contexto de tarea (e.g., Bangert-Drowns et al., 1991; Narciss, 2013; Timmers y Veldkamp, 2011; Timms et al., 2016).

Por último, existe escasa literatura que haya examinado el papel de las diferencias individuales en el procesamiento y la efectividad de la retroalimentación formativa en entornos digitales. Por ejemplo, son necesarias investigaciones que exploren la efectividad de diferentes tipos de EF en función de la competencia lectora o el conocimiento previo de los estudiantes (Gordijn y Nijhof, 2002; Shute, 2008). Igualmente, es posible que diferencias individuales en la velocidad de codificación o el nivel de vocabulario de los estudiantes expliquen el procesamiento de la EF de forma más precisa. Asimismo, sería necesario corroborar el papel moderador de la EF en la relación que existe entre el nivel de conocimiento previo y el rendimiento en las tareas. Siguiendo con las características intrapersonales, es posible que factores relacionados con la habilidad de los estudiantes para refutar conocimientos erróneos, así como las

funciones ejecutivas (e.g., la memoria de trabajo o los procesos inhibitorios), permitan esclarecer diferencias en las operaciones mentales aplicadas cuando se procesa la EF. También es necesario realizar estudios que examinen los factores motivacionales que se relacionan con las decisiones de usar u omitir la EF, ya que algunos estudiantes deciden no procesar esos mensajes ni siquiera cuando responden incorrectamente (e.g., Golke et al., 2015). En definitiva, es necesario realizar más investigaciones para explorar patrones de uso de la EF en función de variables de tarea, contextuales e individuales.

Implicaciones

El presente trabajo tiene implicaciones teóricas y prácticas sobre el uso y procesamiento de la retroalimentación formativa en entornos digitales de aprendizaje, algunas de las cuales ya han sido mencionadas anteriormente. Décadas de investigación confirman que la retroalimentación es un componente instruccional potencialmente eficaz para mejorar el aprendizaje de los estudiantes. A pesar de la amplia variedad de mensajes de retroalimentación formativa que pueden ser proporcionados, la investigación existente ha demostrado que, en términos generales, la EF es la más efectiva para promover el aprendizaje de orden superior (Shute, 2008; Van der Kleij et al., 2015).

Por lo que respecta a las implicaciones teóricas, los resultados indican que los estudiantes emplean la retroalimentación para verificar su modelo de respuesta y confirmar su conocimiento, tal y como postulamos en la primera fase del modelo sobre procesamiento de la retroalimentación. Cuando los estudiantes no están seguros de sus respuestas o cuando éstas son erróneas y son informados de ello, los estudiantes emplean la EF para realizar modificaciones en sus representaciones mentales, como postulamos en la segunda fase del modelo. Estas actualizaciones suelen ser simples y consisten en la codificación de la respuesta correcta cuando se proporciona retroalimentación KCR. Además, observamos que el uso de la EF varía en función de factores externos, como la disponibilidad del texto durante la tarea o la presencia de retroalimentación correctiva KR y KCR. Si la retroalimentación no incluye retroalimentación correctiva, entonces los estudiantes emplean la EF para intentar encontrar pistas que les permitan identificar la respuesta correcta. Los resultados de los dos últimos estudios apoyan esta idea, ya que los estudiantes tienden a acceder proporcionalmente a la misma cantidad de mensajes EF después de acertar y fallar

cuando desconocen la corrección de sus respuestas, lo que sugiere que son imprecisos a la hora de monitorizar la precisión de sus respuestas.

Los resultados también sugieren que es necesario distinguir entre diferentes tipos de retroalimentación. Mientras que la retroalimentación KR es simple y puede ser procesada sin apenas esfuerzo, los mensajes KCR y EF requieren esfuerzo consciente, lo que implica una decisión explícita de procesar su contenido. Mientras que el valor de la retroalimentación KR es simplemente informar al alumno sobre la corrección de sus respuestas, la retroalimentación KCR y EF pueden contribuir al aprendizaje de los estudiantes, ya que reducen la distancia entre el nivel de rendimiento actual y el deseado (Van der Kleij et al., 2015). Por tanto, la retroalimentación KR puede considerarse la puerta de acceso a la EF después de proporcionar respuestas incorrectas. Asimismo, los modelos teóricos sobre el procesamiento de la retroalimentación, como el modelo LP-FILE (Timms et al., 2016), deberían incorporar las operaciones cognitivas y metacognitivas observadas en el segundo estudio, ya que pueden estar relacionadas con la verificación del modelo de respuesta y con el proceso de revisión de conocimientos.

Por lo que respecta a las implicaciones prácticas, hay que señalar que los estudios presentados ofrecen información valiosa para el diseño de entornos digitales, como los libros de texto electrónicos o los sistemas de tutorización inteligentes, en los que la retroalimentación formativa puede facilitar la comprensión y aprendizaje de los estudiantes al resolver tareas como las planteadas en estos estudios. Estos entornos digitales pueden proporcionar tareas y almacenar información sobre la ejecución del estudiante, al objeto de proporcionar retroalimentación inmediata y adaptada. Los estudios presentados emplearon dos tipos de mensajes EF diferentes; si bien uno estaba centrado en la precisión de los estudiantes para responder preguntas de comprensión y localizar información relevante para responder cada pregunta, el otro proporcionaba explicaciones sobre los conceptos evaluados en unas preguntas de aprendizaje de un texto expositivo de física. Los resultados de los dos primeros estudios indican que proporcionar EF específica sobre los segmentos de información relevante para la tarea puede mejorar el procesamiento de la información textual cuando se responden preguntas de comprensión de un texto disponible. Esta información permite al estudiante establecer comparaciones entre su ejecución y el estándar establecido, aprendiendo así a evaluar la relevancia textual durante la propia tarea. De hecho, esta EF es capaz de mejorar tanto la comprensión de los estudiantes como las habilidades para buscar y evaluar la relevancia textual (Llorens et al., 2016). Los tutores inteligentes

de aprendizaje como TuinLEC (Vidal-Abarca et al., 2014) pueden beneficiarse de estos resultados, ya que esta herramienta digital proporciona mensajes EF sobre las estrategias de búsqueda de información que realizan los estudiantes de primaria y secundaria al responder preguntas de comprensión de un texto disponible.

Los dos últimos estudios proporcionan información valiosa para el diseño de libros de texto electrónicos. Estos estudios permiten identificar procedimientos con los que fomentar el uso de la EF al responder preguntas de un texto expositivo. Por ejemplo, parece aconsejable proporcionar retroalimentación KR para facilitar la fase de verificación de respuestas y reducir la discrepancia entre el nivel de seguridad en la respuesta y su precisión. Sin embargo, sería desaconsejable proporcionar KCR si queremos que los estudiantes usen la EF, aunque es posible que esta información sea útil para corregir el modelo de respuesta validado. Al proporcionar retroalimentación KR y KCR podemos asumir que los estudiantes emplearán la EF principalmente después de responder incorrectamente. A pesar de que muchos programas informáticos han sido diseñados para no proporcionar retroalimentación tras responder correctamente, consideramos que los entornos digitales deben proporcionar EF después de acertar, ya que la retroalimentación aumenta la retención de respuestas correctas con bajos niveles de confianza, lo que significa que la retroalimentación es capaz no sólo de corregir errores sino también de corregir imprecisiones metacognitivas para preguntas respondidas correctamente (Butler et al., 2008). Así, los entornos digitales deberían proporcionar EF junto con retroalimentación correctiva (Lefevre y Cox, 2016; Mason y Bruning, 2001), especialmente cuando se promueve el aprendizaje de conocimientos complejos.

En términos generales, los resultados sugieren que los estudiantes monitorizan y autorregulan el uso de la retroalimentación. Como ilustramos en el modelo teórico sobre el procesamiento de la retroalimentación, los estudiantes han de valorar si quieren o necesitan procesar la EF para revisar su conocimiento y generar actualizaciones en su modelo de respuesta. Los estudiantes no procesan la EF automáticamente cuando es recibida, sino que requiere de un proceso activo de toma de decisiones. Así, los entornos digitales deberían ofrecer a los estudiantes la oportunidad de acceder voluntariamente a alguna información específica. Relacionado con lo anterior, los hallazgos también sugieren que los estudiantes pueden no estar dispuestos a invertir mucho tiempo en procesar la retroalimentación, por lo que se debe valorar cuánta información deben contener los mensajes EF e implementar procedimientos que fomenten su uso.

Por otra parte, cabe destacar que existen factores externos e internos que pueden jugar un papel importante en el uso y la efectividad de la EF. Por ejemplo, la disponibilidad de los textos durante la tarea es un factor externo que influye en el uso de la EF. Asimismo, la eficacia de la EF parece ser diferente según el nivel de conocimiento previo del alumnado, como observamos en el cuarto estudio. Los estudiantes con bajos niveles de conocimiento previo en el área de ciencias parecen beneficiarse de la EF si están dispuestos a utilizar esta información a menudo, mientras que los estudiantes con altos niveles de conocimiento previo no se benefician de la misma manera. Es posible que estos últimos necesiten otro tipo de retroalimentación, probablemente menos elaborada.

En definitiva, los estudios presentados en esta tesis abren el camino a nuevas investigaciones centradas en el análisis del procesamiento de la retroalimentación formativa en entornos digitales, así como a las intervenciones destinadas a mejorar la competencia de los estudiantes en el uso y procesamiento de la retroalimentación formativa (Carless y Boud, 2018). No obstante, es necesario llevar a cabo nuevos estudios que examinen el papel de la retroalimentación en otros dominios de conocimiento y con alumnado de otras etapas educativas, evaluando el papel de factores internos y externos que pueden ser determinantes en el uso y la efectividad de la EF. Estos y nuevos resultados permitirán establecer principios generales sobre el uso y procesamiento de la retroalimentación en entornos digitales de aprendizaje.

REFERENCIAS BIBLIOGRÁFICAS ¹

- Afflerbach, P. (1986). The influence of prior knowledge on expert readers' importance assignment processes. En J. A. Niles y R.V. Lalik (Eds.), *Solving problems in literacy: Learners, teachers, and researchers* (pp. 30-40). Chicago, IL: National Reading Conference.
- Aleven, V., Stahl, E., Schworm, S., Fischer, F. y Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73(3), 277–320. doi:10.3102/00346543073003277
- Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10–14. doi:10.3102/0013189X032008010
- Anmarkrud, Ø., McCrudden, M. T., Bråten, I. y 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

¹ Estas referencias corresponden a las citas incluidas en la Introducción y las Conclusiones generales de la presente tesis.

- Artelt, C., Schiefele, U., y Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, 16(3), 363-383. doi: 10.1007/BF03173188
- Attali, Y. y van der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers & Education*, 110, 154-169. doi:10.1016/j.compedu.2017.03.012
- Azevedo, R., y Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13(2), 111–127. doi: 10.2190/9LMD-3U28-3A0G-FTQT
- Azevedo, R. y Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, 96(3), 523–535. doi:10.1037/0022-0663.96.3.523
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A. y Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213–238. doi:10.3102/00346543061002213
- Bjork, R. A., Dunlosky, J. y Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. doi:10.1146/annurev-psych-113011-143823
- Black, P., y Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7-73. doi:10.1080/0969595980050102
- Black, P. y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. doi:10.1007/s11092-008-9068-5

- Bohn-Gettler, C. M. y Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, 39(3), 206-219. doi:10.1016/j.cedpsych.2014.05.003
- Boscolo, P. y Mason, L. (2003). Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts. *The Journal of Experimental Education*, 71(2), 126-148. doi:10.1080/00220970309602060
- Bown, A. (2017). Elaborative feedback to enhance online second language reading comprehension. *English Language Teaching*, 10(12), 164-171. doi:10.5539/elt.v10n12p164
- Bransford, J. (2000). *How people learn: Brain, mind, experience, and school*. Washington, USA: National Academies Press.
- Britt, M. A., Rouet, J.-F. y Durik, A. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. New York, NY: Routledge.
- Brod, G., Werkle-Bergner, M. y Shing, Y. L. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience*, 7, 1-13. doi.org/10.3389/fnbeh.2013.00139
- Burton, C. y Daneman, M. (2007). Compensating for a limited working memory capacity during reading: Evidence from eye movements. *Reading Psychology*, 28(2), 163–186. doi:10.1080/02702710601186407
- Butler, A. C., Godbole, N. y Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology*, 105(2), 290-298. doi:10.1037/a0031026
- Butler, A. C., Karpicke, J. D. y Roediger III, H. L. (2008). Correcting a metacognitive error: Feedback enhances retention of low-confidence correct responses. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 34, 918 –928.
doi:10.1037/0278-7393.34.4.918
- Butler, A. C. y Roediger III, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, 36(3), 604 – 616. doi:10.3758/MC.36.3.604
- Butler, D. L., y Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
doi:10.3102/00346543065003245
- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in higher education*, 31(2), 219-233. doi: 10.1080/03075070600572132
- Carless, D. y Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. doi:10.1080/02602938.2018.1463354
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, 21(5), 279-283. doi:10.1177/0963721412452728
- Cataldo, M. G. y 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(4), 791-799. doi:10.1037/0022-0663.92.4.791
- Cerdán, R., Gilabert, R. y Vidal-Abarca, E. (2011). Selecting information to answer questions: Strategic individual differences when searching texts. *Learning and Individual Differences*, 21(2), 201–205. doi:10.1016/j.lindif.2010.11.007
- Cerdán, R., Gilabert, R. y Vidal-Abarca, E. (2013). Self-generated explanations on the question demands are not always helpful. *The Spanish Journal of Psychology*, 16, e26, 1-11. doi:10.1017/sjp.2013.45

- Cerdán, R. y Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, 100(1), 209–222. doi:10.1037/0022-0663.100.1.209
- Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R. y 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
- Chi, M. T. H., Feltovich, P. J. y Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152. doi:10.1207/s15516709cog0502_2
- Chinn, C. A. y Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63(1), 1-49. doi:10.3102/00346543063001001
- Conole, G. (2016). MOOCs as disruptive technologies: Strategies for enhancing the learner experience and quality of MOOCs. *RED- Revista de Educación a Distancia*, 50. Retrieved from www.um.es/ead/red/50/conole.pdf
- Corbalan, G., Kester, L. y Van Merriënboer, J. J. G. (2009). Dynamic task selection: Effects of feedback and learner control on efficiency and motivation. *Learning and Instruction*, 19, 455–465. doi:10.1016/j.learninstruc.2008.07.002
- Corbalan, G., Paas, F. y Cuypers, H. (2010). Computer-based feedback in linear algebra: Effects on transfer performance and motivation. *Computers & Education*, 55, 692–703. doi:10.1016/j.compedu.2010.03.002
- Crombach, M. J., Boekaerts, M. y Voeten, M. J. M. (2003). Online measurement of appraisals of students faced with curricular tasks. *Educational and Psychological Measurement*, 63(1), 96-111. doi:10.1177/0013164402239319

- Cromley, J. G., Snyder-Hogan, L. E. y Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687-700. doi:10.1037/a0019452
- Diakidoy, I-A. N. (1999). Comprehension and learning from scientific text. En A. Gagatsis (Ed.), *A multidimensional approach to learning in mathematics and science* (pp. 361–389). Thessaloniki, GR: Art of Text.
- Diakidoy, I-A. N. y Kendeou, P. (2001). Facilitating conceptual change in astronomy: A comparison of the effectiveness of two instructional approaches. *Learning & Instruction*, *11*(1), 1-20. doi:10.1016/S0959-4752(00)00011-6
- Dillon, R. F. (1997). Dynamic testing. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 164–186). Westport: Greenwood Press.
- D'Mello, S. K., Craig, S. D., Fike, K. y Graesser, A. C. (2009). Responding to learners' cognitive-affective states with supportive and shakeup dialogues. In J. A. Jacko (Ed.), *Lecture Notes in Computer Science, Vol. 5612: Human-computer interaction—Ambient, ubiquitous and intelligent interaction* (pp. 595–604). Berlin, Germany: Springer-Verlag.
- Dörfler, T., Golke, S. y Artelt, C. (2017). Evaluating prerequisites for the development of a dynamic test of reading competence: Feedback effects on reading comprehension in children. En D. Leutner, J. Fleischer, J. Grünkorn, y E. Klieme (Eds.), *Competence Assessment in Education* (pp. 487-503). Cham, Switzerland: Springer. doi:10.1007/978-3-319-50030-0_28
- Dunlosky, J., Hartwig, M. K., Rawson, K. A. y Lipko-Speed, A. R. (2010). Improving college students' evaluation of text learning using idea-unit standards. *The*

- Quarterly Journal of Experimental Psychology*, 64(3), 467-484.
doi:10.1080/17470218.2010.502239
- Dunlosky, J., y Lipko-Speed, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16(4), 228–232. doi:10.1111/j.1467-8721.2007.00509.x
- Dunlosky, J. y Rawson, K. A. (2015). Do students use testing and feedback while learning? A focus on key concept definitions and learning to criterion. *Learning and Instruction*, 39, 32-44. doi:10.1016/j.learninstruc.2015.05.003
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. y Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4-58. doi:10.1177/1529100612453266
- Durkin, D. (1978–1979). What classroom observation reveals about reading comprehension instruction. *Reading Research Quarterly*, 14(4), 481–533.
- Dweck, C. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- European Commission (2013). The International Computer and Information Literacy Studies (ICILS). Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2014/e-c-icils_en.pdf
- Farr, R., Pritchard, R. y Smitten, B. (1990). A description of what happens when an examinee takes a multiple-choice reading comprehension test. *Journal of Educational Measurement*, 27(3), 209-226. doi:10.1111/j.1745-3984.1990.tb00744.x

- Ferrer, A., Vidal-Abarca, E., Serrano, M. A. y Gilabert, R. (2017). Impact of text availability and question format on reading comprehension processes. *Contemporary Educational Psychology*, 51, 404-415. doi:10.1016/j.cedpsych.2017.10.002
- Finn, B., Thomas, R. y Rawson, K. A. (2018). Learning more from feedback: Elaborating feedback with examples enhances concept learning. *Learning and Instruction*, 54, 104-113. doi:10.1016/j.learninstruc.2017.08.007
- Fox, J.-P., Klein Entink, R. y Timmers, C. (2014). The joint multivariate modeling of multiple mixed response sources: Relating student performances with feedback behavior. *Multivariate Behavioral Research*, 49(1), 54-66. doi:10.1080/00273171.2013.843441
- Fyfe, E. R. (2016). Providing feedback on computer-based algebra homework in middle-school classrooms. *Computers in Human Behavior*, 63, 568-574. doi:10.1016/j.chb.2016.05.082
- Fyfe, E. R., Rittle-Johnson, B. y DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104(4), 1094–1108. doi:10.1037/a0028389
- Gil, L., Martínez, T. y Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education*, 82, 50-59. doi:10.1016/j.compedu.2014.10.026
- Glenberg, A. M. y Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 702-718.
- Goldin, I. M., Koedinger, K. R. y Aleven, V. (2012). Learner differences in hint processing. In K. Yacef, O. Zaïane, H. HersHKovitz, M. Yudelson y J. Stamper

- (Eds.), *Proceedings of the 5th international conference on educational data mining* (pp. 73–80). Worcester, MA: International Educational Data Mining Society.
- Golding, J. M., Graesser, A. C. y Millis, K. K. (1990). What makes a good answer to a question?: Testing a psychological model of question answering in the context of narrative text. *Discourse Processes*, *13*(3), 305-325. doi:10.1080/01638539009544761
- Golke, S., Dörfler, T. y 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
- Gordijn, J. y Nijhof, W. J. (2002). Effects of complex feedback on computer-assisted modular instruction. *Computers & Education*, *39*(2), 183–200. doi:10.1016/S0360-1315(02)00025-8
- Graesser, A. C. (2016). Conversations with AutoTutor help students learn. *International Journal of Artificial Intelligence in Education*, *26*(1), 124-132. doi:10.1007/s40593-015-0086-4
- Graesser, A. C. y Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension. *Advances in Psychology*, *29*, 53-94. doi:10.1016/S0166-4115(08)62732-6
- Graesser, A. C. y Franklin, S. P. (1990). QUEST: A cognitive model of question answering. *Discourse Processes*, *13*(3), 279–303. doi:10.1080/01638539009544760
- Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H., Ventura, M., Olney, A. y Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, *36*, 180–193. doi:10.3758/BF03195563

- Graesser, A. C., McNamara, D. S., Louwerse, M. M. y Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. doi:10.3758/BF03195564
- Graesser, A. C. y Murachver, T. (1985). Symbolic procedures of question answering. In A. C. Graesser y J. B. Black (Eds.), *The psychology of questions* (pp. 15-88). Hillsdale, NJ: Erlbaum.
- Graesser, A. C., Robertson, S. P., Lovelace, E. R. y Swinehart, D. M. (1980). Answers to why-questions expose the organization of story plot and predict recall of actions. *Journal of Verbal Learning and Verbal Behavior*, 19(1), 110-119. doi:10.1016/S0022-5371(80)90572-1
- Graesser, A. C., Singer, M., y Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395. doi:10.1037//0033-295X.101.3.371
- Hacker, D. J. (1998). Self-regulated comprehension during normal reading. En D. J. Hacker, A.C. Graesser, y J. Dunlosky (Eds.), *Metacognition in educational theory and practice* (pp. 165-191). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Hannafin, M. F., Hannafin, K. M., y Dalton, D. W. (1993). Feedback and emerging instructional technologies. En J. V. Dempsey, y G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 263–286). Englewood Cliffs, NJ: Educational Technology.
- Hannon, B. y Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse Processes*, 37(3), 187–204. doi:10.1207/s15326950dp3703_1
- Hattie, J. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Abingdon: Routledge.

- Hattie, J., y Gan, M. (2011). Instruction based on feedback. En R. E. Mayer y P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249-271). New York, NY: Routledge.
- Hattie, J. y Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Higgs, K., Magliano, J. P., Vidal-Abarca, E., Martínez, T. y McNamara, D. S. (2017). Bridging skill and task-oriented reading. *Discourse Processes*, 54(1), 19-39. doi:10.1080/0163853X.2015.1100572
- Jaehnig, W. y Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, 57, 219–232. doi:10.1007/BF03395573
- Johnson, W. L. y Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. *International Journal of Artificial Intelligence in Education*, 26(1), 25–36. doi:10.1007/s40593-015-0065-9
- Johnson, W. L., Rickel, J. W. y Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11(1), 47–78.
- Kaakinen, J. K., Hyönä, J. y Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(3), 447–457. doi:10.1037/0278-7393.29.3.447
- Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, 21(3), 157-163. doi:10.1177/0963721412443552

- Karpicke, J. D. y Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 704-719. doi:10.1037/0278-7393.33.4.704
- Kendeou, P. y O'Brien, E. J. (2014). The Knowledge Revision Components (KReC) framework: Processes and mechanisms. In D. N. Rapp y J. L. G. Braasch (Eds.). *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences*. Cambridge, MA: MIT Press.
- Kendeou, P. y O'Brien, E. J. (2015). Prior knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Text and context* (pp. 151-163). New York, NY: Routledge.
- Kendeou, P. y O'Brien, E. J. (2018). Theories of text processing: A view from the top-down. In M. Schober, D. N. Rapp y M. A. Britt (Eds.). *Handbook of discourse processes (2nd edition)* (pp. 19-33). New York: Routledge Publishing.
- Kendeou, P., Rapp, D. N. y van den Broek, P. (2004). The influence of readers' prior knowledge on text comprehension and learning from text. In R. Nata (Ed.), *Progress in education* (Vol. 13, pp. 189-209). New York: Nova Science.
- Kendeou, P. y van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, 97(2), 235-245. doi:10.1037/0022-0663.97.2.235
- Kendeou, P. y van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567-1577. doi:10.3758/BF03193491
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

- Kintsch, W. y van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394. doi:10.1037/0033-295X.85.5.363
- Kluger, A. N. y Adler, S. (1993). Person- versus computer-mediated feedback. *Computers in Human Behavior*, 9(1), 1-16. doi:10.1016/0747-5632(93)90017-M
- Kluger, A. N., Adler, S. y Fay, C. (1992). Computerized feedback effects on feedback seeking, performance and motivation. In M. I. Nurminen y G. R. S. Weir (Eds.), *Human jobs and computer interfaces* (pp. 131-145). Amsterdam: North-Holland.
- Kluger, A. N. y 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
- Krause, U.-M., Stark, R. y Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, 19(2), 158 –170. doi:10.1016/j.learninstruc.2008.03.003
- Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, 47(2), 211-232. doi:10.3102/00346543047002211
- Kulhavy, R. W., y Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4), 279-308. doi:10.1007/BF01320096
- Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L., y Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, 10(3), 285–291. doi:10.1016/0361-476X(85)90025-6

- Kupiainen, S., Vainikainen, M.-P., Marjanen, J. y Hautamäki, J. (2014). The role of time on task in computer-based low-stakes assessment of cross-curricular skills. *Journal of Educational Psychology, 106*(3), 627-638. doi:10.1037/a0035507
- Langer, P. (2011). The use of feedback in education: A complex instructional strategy. *Psychological Reports, 109*(3), 775-784. doi:10.2466/11.PR0.109.6.775-784
- Langer, P., y Keenan, V. (2000). Text processing and feedback: What have we learned? *Psychological Reports, 87*(2), 406–412. doi:10.2466/pr0.2000.87.2.406
- Langer, P., Keenan, V., y Medosch-Schonbeck, C. M. (1985). A text synthesis approach to feedback. *Psychological Reports, 57*(2), 599–610. doi:10.2466/pr0.1985.57.2.599
- Lee, H. W., Lim, K. Y. 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.
- Lee, H. W., Lim, K. Y. y Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research and Development, 58*(6), 629–648. doi:10.1007/s11423-010-9153-6
- Lefevre, D., y Cox, B. (2016). Feedback in technology-based instruction: Learner preferences. *British Journal of Educational Technology, 47*(2), 248-256. doi:10.1111/bjet.12227
- Lewis, M. R. y Mensink, M. C. (2012). Prereading questions and online text comprehension. *Discourse Processes, 49*(5), 367–390. doi:10.1080/0163853X.2012.662801
- Lipko-Speed, A. R., Dunlosky, J., Hartwig, M., Rawson, K. A., Swan, K., y Cook, D. L. (2009). Using standards to improve middle-school students' accuracy at evaluating

- the quality of their recall. *Journal of Experimental Psychology: Applied*, 15, 307–318. doi:10.1037/a0017599
- Lipko-Speed, A., Dunlosky, J. y Rawson, K. A. (2014). Does testing with feedback help grade-school children learn key concepts in science? *Journal of Applied Research in Memory and Cognition*, 3(3), 171-176. doi:10.1016/j.jarmac.2014.04.002
- Llorens, A. C., Cerdán, R. y Vidal-Abarca, E. (2014). Adaptive formative feedback to improve strategic search decisions in task-oriented reading. *Journal of Computer Assisted Learning*, 30(3), 233–251. doi:10.1111/jcal.12050
- Llorens, A. C., Vidal-Abarca, E. y 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. y 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
- Maier, U., Wolf, N. y Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85-98. doi:10.1016/j.compedu.2015.12.002
- Mañá, A., Vidal-Abarca, E., Domínguez, C., Gil, L., y Cerdán, R. (2009). Papel de los procesos metacognitivos en una tarea de pregunta-respuesta con textos escritos. *Infancia y Aprendizaje*, 32(4), 553–565. doi: 10.1174/021037009789610412
- Mañá, A., Vidal-Abarca, E. y Salmerón, L. (2017). Effect of delay on search decisions in a task-oriented reading environment. *Metacognition and Learning*, 12(1), 113–130. doi:10.1007/s11409-016-9162-x

- Martínez, T., Vidal-Abarca, E., Gil, L. y Gilabert, R. (2009). On-line assessment of comprehension processes. *The Spanish Journal of Psychology*, *12*(1), 308-319. doi:10.1017/S1138741600001700
- Martínez, T., Vidal-Abarca, E., Sellés, P. y Gilabert, R. (2008). Evaluation of comprehension strategies and processes: Test of Comprehension Processes (TCP). *Infancia y Aprendizaje*, *31*(3), 319-332. doi:10.1174/021037008785702956
- Mason, B. y 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
- Mason, L., Tornatora, M. C. y Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, *60*(1), 95-109. doi:10.1016/j.compedu.2012.07.011
- McCrudden, M. T., Magliano, J. P. y Schraw, G. (2011). Relevance in text comprehension. In M. T. McCrudden, J. P. Magliano y G. Schraw (Eds.), *Text relevance and learning from text* (pp. 1-17). Greenwich, CT: Information Age Publishing.
- McCrudden, M. T. y Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*, 113-139. doi:10.1007/s10648-006-9010-7
- McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*(1), 51-62. doi:10.1037/h0087352

- McNamara, D. S. y Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22(3), 247-288. doi:10.1080/01638539609544975
- McNamara, D. S., Kintsch, E., Songer, N. B. y Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition & Instruction*, 14(1), 1-43. doi:10.1207/s1532690xci1401_1
- McNamara, D. S., y Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, 51, 297-384. doi:10.1016/S0079-7421(09)51009-2
- Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, 131(3), 349-363. doi:10.1037//0096-3445.131.3.349
- Metcalfe, J., y Finn, B. (2008) Evidence that judgements of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15(1), 174-179. doi:10.3758/PBR.15.1.174
- Metcalfe, J. y Kornell, N. (2007). Principles of cognitive science in education: The effects of generation, errors, and feedback. *Psychonomic Bulletin & Review*, 14, 225-229. doi:10.3758/BF03194056
- Metcalfe, J., Kornell, N. y Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, 37, 1077-1087. doi:10.3758/MC.37.8.1077
- Metcalfe, J., Kornell, N. y Son, L. K. (2007). A cognitive science based programme to enhance study efficacy in a high and low risk setting. *European Journal of Cognitive Psychology*, 19, 743-768. doi:10.1080/09541440701326063

- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science*, 32, 99–113. doi:10.1023/B:TRUC.0000021811.66966.1d
- Mory, E. H. (2004). Feedback research revisited. En D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Erlbaum.
- Murphy, P. (2007). Reading comprehension exercises online: The effects of feedback, proficiency and interaction. *Language Learning and Technology*, 11(3), 107–129.
- Murphy, P. (2010). Web-based collaborative reading exercises for learners in remote locations: The effects of computer-mediated feedback and interaction via computer-mediated communication. *ReCALL*, 22(2), 112–134. doi:10.1017/S0958344010000030
- Narciss, S. (2004). The impact of informative tutoring feedback and self-efficacy on motivation and achievement in concept learning. *Experimental Psychology*, 51(3), 214–228. doi:10.1027/1618-3169.51.3.214
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. van Merriënboer y M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 125–144). New York, NY: Erlbaum.
- Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. *Digital Education Review*, 23, 7-26.
- Narciss, S. y Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner y R. Brünken (Eds.), *Instructional design for multimedia learning* (pp. 181-195). Münster: Waxmann.

- Narciss, S., Körndle, H., Reimann, G. y Müller, C. (2004). Feedback-seeking and feedback efficiency in web-based learning – How do they relate to task and learner characteristics? In P. Gerjets, P.A. Kirschner, J. Elen y R. Joiner (Eds.), *Instructional design for effective and enjoyable computer- supported learning. Proceedings of the first joint meeting of the EARLI SIGs Instructional Design and Learning and Instruction with Computers* (pp. 377-388). Tuebingen: Knowledge Media Research Center.
- Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Gogvadze, G. y Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, 71, 56-76. doi:10.1016/j.compedu.2013.09.011
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction: Reports of the subgroups*. Washington DC: National Institute of Child Health and Human Development.
- Ness, M. (2011). Explicit reading comprehension instruction in elementary classrooms: Teacher use of reading comprehension strategies. *Journal of Research in Childhood Education*, 25(1), 98–117. doi:10.1080/02568543.2010.531076
- Nicol, D. J., y Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi: 10.1080/03075070600572090
- Nuthall, G. (2007). *The hidden lives of learners*. New Zealand: New Zealand Council for Educational Research.

- Nye, B. D., Graesser, A. C. y Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, 24(4), 427-469. doi:10.1007/s40593-014-0029-5
- OECD (2010). *PISA 2009. Assessment Framework: Key Competencies in Reading, Mathematics and Science*. Paris: OECD Publishing.
- OECD (2017). *PISA 2015. Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving (revised edition)*. Paris: OECD Publishing.
- Olney, A. M., Graesser, A. C., y Person, N. K. (2010). Tutorial dialog in natural language. In R. Nkambou, J. Bourdeau, y R. Mizoguchi (Eds.), *Advances in intelligent tutoring systems: Studies in computational intelligence (pp. 181-206)*. Berlin: Springer.
- Otero, J. (2002). Noticing and fixing difficulties in understanding science texts. En J. Otero, J. A. León, A. Graesser (Eds.). *The psychology of science text comprehension* (pp. 281–307). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Ozuru, Y., Best, R., Bell, C., Witherspoon, A. y McNamara, D. S. (2007). Influence of question format and text availability on the assessment of expository text comprehension. *Cognition and Instruction*, 25(4), 399–438. doi:10.1080/07370000701632371
- Ozuru, Y., Dempsey, K. y McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, 19(3), 228-242. doi:10.1016/j.learninstruc.2008.04.003
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341. doi:10.1007/s10648-006-9029-9

- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich y M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press.
- Pintrich, P. R. y Zusho, A. (2002). Student motivation and self-regulated learning in the college classroom. In J. C. Smart y W.G. Tierney (Eds), *Higher Education: Handbook of theory and research* (vol. XVII). Dordrecht: Springer. doi:10.1007/978-94-010-0245-5
- Pressey, S. L. (1926). A simple device which gives tests and scores- and teaches. *School and Society*, 23, 373-376.
- Pressey, S. L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *The Journal of Psychology*, 29(2), 417-447. doi:10.1080/00223980.1950.9916043
- Pressley, M., Wharton-McDonald, R., Mistretta-Hampston, J. y Echevarria, M. (1998). Literacy instruction in 10 fourth-grade classrooms in upstate New York. *Scientific Studies of Reading*, 2(2), 159-194. doi:10.1207/s1532799xssr0202_4
- Pridemore, D. R. y Klein, J. D. (1991). Control of feedback in computer-assisted instruction. *Educational Technology Research and Development*, 39(4), 27-32. doi:10.1007/bf02296569
- Pridemore, D. R. y Klein, J. D. (1995). Control of practice and level of feedback in computer-based instruction. *Contemporary Educational Psychology*, 20(4), 444-450. doi:10.1006/ceps.1995.1030
- Rakoczy, K., Harks, B., Klieme, E., Blum, W. y Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction*, 27, 63-73. doi:10.1016/j.learninstruc.2013.03.002

- Ramos, L. y 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
- Rawson, K. A., y Dunlosky, J. (2007). Improving students' self-evaluation of learning for key concepts in textbook materials. *European Journal of Cognitive Psychology*, 19(4-5), 559–579. doi:10.1080/09541440701326022
- Roelle, J. y Berthold. K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, 49(2), 142-156. doi:10.1016/j.learninstruc.2017.01.008
- Rouet, J. F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Erlbaum.
- Rouet, J-F. y Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano y G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Greenwich, CT: Information Age Publishing.
- Rouet, J-F., Britt, M. A. y Durik, A. M., (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200-215. doi:10.1080/00461520.2017.1329015
- Rouet, J-F. y Coutelet, B. (2008). The acquisition of document search strategies in grade school students. *Applied Cognitive Psychology*, 22(3), 389–406. doi:10.1002/acp.1415
- Rupp, A. A., Ferne, T. y Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, 23(4), 441–474. doi:10.1191/0265532206lt337oa

- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144. doi:10.1007/BF00117714
- Salmerón, L., Vidal-Abarca, E., Martínez, T., Mañá, A., Gil, L. y Naumann, J. (2015). Strategic decisions in task-oriented reading. *The Spanish Journal of Psychology*, 18, 1–10. doi:10.1017/sjp.2015.101
- Sánchez, E., y García, J. R. (2015). Understanding teachers as learners in reading comprehension mentoring. En H. Tillema, G. J. van der Westhuizen, y K. Smith (Eds.), *Mentoring for learning* (pp. 227–255). Rotterdam, The Netherlands: SensePublishers.
- Sánchez, E., García, J. R. y Rosales, J. (2010). *La lectura de los textos en el aula: Qué hacen los alumnos y sus profesores cuando usan los textos y qué se podría llegar a hacer*. Barcelona: Graó.
- Schaffner, E. y Schiefele, U. (2013). The prediction of reading comprehension by cognitive and motivational factors: Does text accessibility during comprehension testing make a difference? *Learning and Individual Differences*, 26, 42–54. doi:10.1016/j.lindif.2013.04.003
- Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology*, 103(4), 877-896. doi:10.1037/a0023731
- Schunk, D. H. y Rice, J. M. (1991). Learning goals and progress feedback during reading comprehension instruction. *Journal of Literacy Research*, 23(3), 351-364. doi:10.1080/10862969109547746
- Schunk, D. H. y Rice, J. M. (1993). Strategy fading and progress feedback: Effects on self-efficacy and comprehension among students receiving remedial reading

- services. *The Journal of Special Education*, 27(3), 257-276.
<http://dx.doi.org/10.1177/002246699302700301>
- Schutz, P. A. (1993). Additional influences on response certitude and feedback requests. *Contemporary Educational Psychology*, 18(4), 427-441.
doi:10.1006/ceps.1993.1032
- Schwartz, D. L., Tsang, J. M. y Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. New York, NY: WW Norton & Company.
- Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, 41(1), 159-189. doi:10.3102/00028312041001159
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. doi:10.3102/0034654307313795
- Shute, V. J. y Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33(1), 1-19. doi:10.1111/jcal.12172
- Skinner, B.F. (1958). Teaching machines. *Science*, 128(3330), 969-977.
- Smits, M., Boon, J., Sluijsmans, D. M. A. y Van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: Effects on learning as a function of prior knowledge. *Interactive Learning Environments*, 16(2), 183–193.
doi:10.1080/10494820701365952
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

- Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences, 20*(5), 415-420. doi:10.1016/j.lindif.2010.04.002
- Tärning, B. (2018). Review of feedback in digital applications-Does the feedback they provide support learning? *Journal of Information Technology Education: Research, 17*, 247-283. doi:10.28945/4104
- Thiede, K. W., Anderson, M. C. M., y Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning from texts. *Journal of Educational Psychology, 95*(1), 66–73. doi:10.1037/0022-0663.95.1.66
- Thiede, K. W., y Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 1024–1037. doi:10.1037/0278-7393.25.4.1024
- Thiede, K. W., Griffin, T. D., Wiley, J., y Redford, J. (2009). Metacognitive monitoring during and after reading. En D. J. Hacker, J. Dunlosky y A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 85–106). New York: Routledge.
- Timmers, C. F., Braber-van den Broek, J. y van den Berg, S. M. (2013). Motivational beliefs, student effort, and feedback behaviour in computer-based formative assessment. *Computers & Education, 60*(1), 25-31. doi:10.1016/j.compedu.2012.07.007
- Timmers, C. F. y Veldkamp, B. P. (2011). Attention paid to feedback provided by a computer-based assessment for learning on information literacy. *Computers & Education, 56*(3), 923–930. doi:10.1016/j.compedu.2010.11.007

- Timms, M., DeVelle, S. y Lay, D. (2016). Towards a model of how learners process feedback: A deeper look at learning. *Australian Journal of Education*, 60(2), 128-145. doi:10.1177/0004944116652912
- van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, 328(5977), 453–456. doi:10.1126/science.1182594
- Van der Kleij, F. M., Eggen, T. J. H. M., Timmers, C. F. y Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, 58(1), 263–272. doi:10.1016/j.compedu.2011.07.020
- Van der Kleij, F. M., Feskens, R. C. W. y 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
- Van der Kleij, F. M., Timmers, C. F. y Eggen, T. J. H. M. (2011). The effectiveness of methods for providing written feedback through a computer-based assessment for learning: A systematic review. *CADMO*, 19, 21–39. doi:10.3280/CAD2011-001004
- van Dijk, T. y Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vidal-Abarca, E., Gilabert, R., Ferrer, A., Ávila, V., Martínez, T., Mañá, A., ... y Serrano, M. Á. (2014). TuinLEC, an intelligent tutoring system to improve reading literacy skills/TuinLEC, un tutor inteligente para mejorar la competencia lectora. *Infancia y Aprendizaje*, 37(1), 25–56. doi:10.1080/02103702.2014.881657
- Vidal-Abarca, E., Mañá, A. y 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., Gil, L., García, A., y Máñez, I. (2019). Learning tasks in electronic environments: Advances towards interactive eTextbooks. En K. Millis, D. L. Long, J. P. Magliano, y K. Wiemer (Eds.), *Deep comprehension: Multi-disciplinary approaches to understanding, enhancing, and measuring comprehension*. New York, NY: Routledge.
- Vidal-Abarca, E., Salmerón, L., y Mañá, A. (2011). Individual differences in task-oriented reading. En M. T. McCrudden, J. P. Magliano, y G. Schraw (Eds.), *Relevance instructions and goal-focusing in text learning*. Greenwich, CT: Information Age Publishing.
- Waddick, J. (1994). Case study: The creation of a computer learning environment as an alternative to traditional lecturing methods in chemistry. *Educational & Training Technology International*, 31(2), 98-103. doi:10.1080/0954730940310203
- Winne, P. H. (2004). Students' calibration of knowledge and learning processes: Implications for designing powerful software learning environments. *International Journal of Educational Research*, 41(6), 466-488. doi:10.1016/j.ijer.2005.08.012
- Winne, P. H., y Hadwin, A. (1998). Studying as self-regulated learning. En D. J. Hacker, J. Dunlosky, y A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.

ANEXOS

ANEXO A: TEXTOS Y PREGUNTAS DE LOS ESTUDIOS 1 Y 2

ANEXO B: TEXTO Y PREGUNTAS DEL ESTUDIO 3

ANEXO C: TEXTO Y PREGUNTAS DEL ESTUDIO 4

ANEXO A

TEXTOS Y PREGUNTAS DE LOS ESTUDIOS 1 Y 2

TEXTO 1.

EL PINGÜINO

El pingüino es un ave que se cuenta por millones y que no conoce el hemisferio Norte. Vive en lugares muy diversos. Sólo algunas especies se han adaptado a lugares cálidos como las islas Galápagos, pero la gran mayoría de los pingüinos prefiere las regiones frías de Sudamérica, y de continentes como África, Australia y sobre todo, la Antártida, que es su hábitat más común ya que no hay sol la mayor parte del año. Los científicos han descrito 17 especies. Los más pequeños son los pingüinos azules que viven en Australia y miden 40 centímetros. La especie más grande es el elegante pingüino Emperador, enigmático habitante de la eterna noche invernal del Polo Sur, que llega a medir hasta 130 centímetros y puede pesar más de 30 kilos.

En lo que coinciden todos los pingüinos es en el recubrimiento de su cuerpo. Tienen una espesa capa de grasa y sobre ella un abrigo de plumas cortas y muy densas colocadas de tal manera que forman cámaras de aire aislantes del frío ambiente. También coinciden en su solidaridad, es decir, se ayudan unos a otros, incluso sin ser de la misma familia, lo que les permite hacer frente a los paisajes hostiles y climas duros en los que viven.

Pero lo más enternecedor y original es su comportamiento reproductor. Los pingüinos Adelia, una de las especies de pingüinos antárticos, incuban sus huevos y crían a sus pequeños turnándose el macho y la hembra. Suelen formar parejas estables con bajísimo número de divorcios. Mientras uno se queda con el huevo, el otro progenitor se aleja hasta el agua para buscar comida. Después del nacimiento, ambos padres siguen compartiendo el cuidado de los polluelos y la búsqueda de comida. Comen peces y krill, un crustáceo parecido a diminutas gambas muy abundante en las aguas del Polo Sur y que sirve también de alimento para las ballenas.

Sin embargo, en el caso del pingüino Emperador, otra de las especies antárticas, es el padre el que asume toda la incubación del huevo durante nueve semanas. En ese tiempo, la madre se marcha a la costa en busca de algo para comer. No es un viaje de placer, tendrá que poner los cinco sentidos en sus travesías para no ser devorada por alguno de sus peores enemigos. En el agua, el peligro mayor son las orcas; en el borde del casquete polar también acechan las focas Leopardo, animales grandes y solitarios, con cara de pocos amigos. Ambos disfrutan enormemente zampándose pingüinos. Mientras la madre está fuera, el padre sobrevive echando mano de sus michelines, o reservas de grasa.

Una vez que nace el polluelo del pingüino Emperador, viene a darle el relevo su esposa, que asume la cría del pequeño durante seis semanas. Durante ese tiempo, el macho emprende largas y pesadísimas caminatas de hasta 160 kilómetros en busca de la gran mariscada que le reponga de la paternidad; cuidar el huevo le ha supuesto perder hasta un tercio de su peso corporal. Cuando el pequeño ya ha cumplido alrededor de 7 semanas, lo llevan a las guarderías que se montan en las enormes comunidades de pingüinos del Polo Sur, bajo la atenta vigilancia de unos pocos adultos responsables. De esta forma ambos miembros de la pareja pueden marcharse en busca de comida.

PREGUNTAS TEXTO 1: EL PINGÜINO

1. Entre los pingüinos Emperador la incubación del huevo la realiza:
 - a) El pingüino Emperador macho.
 - a) Primero el macho y luego la hembra, indistintamente.
 - b) El macho o la hembra.
 - c) Primero la hembra y luego el macho.

2. ¿Qué forma cámaras de aire en los pingüinos?
 - a) Las plumas que recubren su cuerpo.
 - b) Una capa de grasa que recubre el cuerpo.
 - c) Un abrigo de pelo que tapa su cuerpo.
 - d) Una piel especial que recubre su cuerpo.

3. El krill es:
 - a) Un pequeño pez que sirve de alimento a pingüinos y ballenas.
 - b) Un crustáceo que se alimenta de peces y vive en el Polo Sur.
 - c) Una especie de gamba que sirve de alimento a los pingüinos.
 - d) Un crustáceo que se alimenta de gambas, al igual que los pingüinos.

4. Los pingüinos Emperador y los pingüinos Adelia se diferencian en:
 - a) La forma de organizar el cuidado de las crías.
 - b) El lugar donde viven.
 - c) El recubrimiento de su cuerpo para aislarse del frío.
 - d) La clase de alimentos que comen.

5. ¿Qué tienen en común todos los pingüinos?
 - a) El continente donde viven y el recubrimiento de su cuerpo.
 - b) El recubrimiento y el tamaño de su cuerpo.
 - c) La ayuda que se prestan y el tamaño de su cuerpo.
 - d) El recubrimiento de su cuerpo y la ayuda que se prestan.

6. ¿Por qué el macho y la hembra de pingüinos Emperador pueden marcharse juntos en busca de comida?
 - a) Porque el pequeño pingüino ya puede vivir solo.
 - b) Porque hay pingüinos que habitan en el Polo Sur.
 - c) Porque dejan a la cría escondida.
 - d) Porque otros pingüinos adultos cuidan de sus hijos.

7. Los pingüinos Adelia tienen un comportamiento reproductor original porque:
 - a) Los pingüinos Adelia crían e incuban sus huevos.
 - b) Los padres se ayudan para incubar y cuidar a las crías.
 - c) Son los machos Adelia los que incuban los huevos.
 - d) Cambian de pareja mientras cuidan de los huevos.

8. ¿Qué hacen los pingüinos Emperador nada más nacer la cría?
 - a) Los padres cuidan juntos del pequeño durante varias semanas.
 - b) El macho cuida a la cría y la hembra busca comida.
 - c) Dejan a la cría al cuidado de adultos responsables y van a buscar comida.
 - d) La hembra cuida a la cría mientras el macho se va a buscar comida.

9. Un ejemplo de la solidaridad de los pingüinos Emperador es que:
- a) El pingüino Emperador macho incuba los huevos.
 - b) El macho y la hembra se turnan para cuidar a sus hijos.
 - c) Pingüinos adultos cuidan de grupos de pingüinos pequeños.
 - d) Pueden hacer frente a paisajes hostiles y climas muy duros.
10. ¿Qué animales se comen a los pingüinos?
- a) Orcas y focas.
 - b) Ballenas y orcas.
 - c) Osos y ballenas.
 - d) Focas y osos.

TEXTO 2.

LOS SIOUX

Hace más de doscientos años, en la mayor parte del territorio de América del Norte habitaban los sioux en campamentos de tiendas, igual que otros indios de las praderas. Vivían pacíficamente y su principal riqueza eran los bisontes, de los que obtenían carne para alimentarse, pieles para abrigarse y huesos con los que fabricaban utensilios de uso cotidiano.

La historia de los sioux está llena de guerras y conflictos. Las batallas más famosas tuvieron lugar en la segunda mitad del siglo XIX. Varios hechos hostiles de los colonos y el Gobierno estadounidense provocaron conflictos sangrientos. Los colonos blancos buscando tierras y los mineros en busca de oro iniciaron una continua invasión de los territorios indios, matando muchos rebaños de bisontes. El gobierno intentó encerrar a los sioux en reservas. Todo ello fue provocando actos feroces y salvajes por parte de los indios. Era un intento de recuperar su propia tierra y su libertad.

En 1863, las tropas de ejército expulsaron a todas las tribus que vivían en Minnesota, la tierra de sus antepasados. Dos años más tarde el ejército estableció una línea fortificada a lo largo del río Missouri que mantenía a los sioux alejados de las grandes manadas de bisontes y de las minas de oro. Grupos de guerreros sioux acosaron durante tres años las zonas fortificadas por el ejército.

En 1868, el Gobierno se vio obligado a firmar la paz y emprendió una política de ayuda económica y cultural a los indios que aceptaron establecerse en las reservas. Pero la paciencia de los indios se acabó cuando las tropas del general Custer penetraron en territorio sioux para proteger a los mineros llegados en busca de oro. En 1876 un grupo de indios tendió una emboscada mortal a las tropas de Custer cuando éstas se disponían a atacar un campamento. Ningún blanco sobrevivió. Esta sangrienta batalla produjo una fuerte reacción del Gobierno. Los sioux sufrieron nuevas reducciones en su territorio, y sus condiciones de vida empeoraron.

La última resistencia armada de los sioux tuvo lugar en 1890, y estuvo ligada a la aparición de un movimiento religioso llamado “La Danza del Espíritu”. El movimiento prometía la vuelta de los rebaños de bisontes y la expulsión de los colonos. Esta creencia se extendió de una tribu a otra y se reanudaron las luchas sangrientas contra los blancos. El movimiento suscitó verdadero terror entre los colonos, que pidieron la intervención del Gobierno.

En diciembre de 1890 un grupo de indios rebeldes cayó en una emboscada y toda la tribu fue exterminada, incluidos ancianos, mujeres y niños. Esta atrocidad significó el fin de las revueltas sioux. En ese momento los indios se vieron forzados a aceptar las condiciones del Gobierno americano. En la actualidad, la mayoría de los sioux viven pobremente en las reservas de Dakota del Norte y del Sur.

PREGUNTAS TEXTO 2: LOS SIOUX

1. ¿Qué mantenía alejados a los sioux de las manadas de bisontes?
 - a) Minnesota, la tierra de sus antepasados.
 - b) Las minas de oro de los colonos.
 - c) El cauce del río Missouri.
 - d) Una línea fortificada del ejército.

2. Hasta hace más de 200 años los sioux vivían pacíficamente porque:
 - a) Convivían junto a otros indios de las praderas.
 - b) Habitaban en campamentos de tiendas situados en las praderas.
 - c) Sus tierras todavía no habían sido invadidas por los blancos.
 - d) Había grandes manadas de bisontes, de los que obtenían alimento.

3. En resumen, la causa de las guerras entre los indios y el gobierno americano desde 1868 hasta 1876 fue que:
 - a) El gobierno firmó la paz y emprendió una política de ayuda a los indios.
 - b) En 1876 un grupo de indios tendió una emboscada a las tropas de Custer.
 - c) El gobierno no cumplió los acuerdos que había firmado con los indios.
 - d) Las condiciones de vida de los sioux empeoraron tras la guerra con el ejército.

4. Los indios en un intento de recuperar su libertad:
 - a) Pactaron con el gobierno para vivir en las reservas.
 - b) Fueron encerrados en reservas por el gobierno.
 - c) Cometieron actos feroces contra los colonos.
 - d) Fueron expulsados de Minnesota en 1863.

5. Las luchas de los sioux terminaron definitivamente cuando:
 - a) Apareció un movimiento religioso llamado “La Danza del Espíritu”.
 - b) Los indios rebeldes aceptaron vivir en Dakota del Norte y del Sur.
 - c) Los sioux cometieron una atrocidad en una revuelta.
 - d) Una tribu de indios rebeldes fue exterminada.

6. A lo largo de la historia, ¿por qué se enfrentaron los sioux a los blancos?
 - a) Porque los blancos les quitaban las tierras a los indios.
 - b) Porque el ejército americano asesinó a una tribu de indios.
 - c) Porque apareció el movimiento religioso “La Danza del Espíritu”.
 - d) Porque los sioux derrotaron a las tropas de Custer.

7. ¿De qué vivían principalmente los sioux?
 - a) De utensilios de uso cotidiano.
 - b) De los bisontes.
 - c) De pieles y huesos.
 - d) De campamentos de tiendas.

8. ¿Con qué estuvieron relacionadas las últimas rebeliones sioux?
 - a) Con un movimiento religioso indio.
 - b) Con las luchas sangrientas entre los blancos.
 - c) Con una emboscada mortal a las tropas de Custer.
 - d) Con unos espíritus religiosos que danzaban.

9. El gobierno intentó inicialmente encerrar a los sioux en las reservas para que:

- a) Los colonos y los mineros ocuparan las tierras indias.
- b) Los indios recuperaran su tierra y su libertad.
- c) Muchos rebaños de bisontes fueran eliminados.
- d) Los indios provocaran actos salvajes.

10. ¿Quién no sobrevivió en 1876?

- a) Un grupo de indios.
- b) Un grupo de mineros blancos.
- c) Todo un campamento indio.
- d) Los soldados de Custer.

ANEXO B

TEXTO Y PREGUNTAS DEL ESTUDIO 3

PRE-TEST. CUESTIONARIO DE CONOCIMIENTO PREVIO

1	Dos objetos con el mismo volumen tienen la misma densidad.	V	F	NS
2	El aire de la atmósfera no pesa.	V	F	NS
3	La temperatura es una propiedad de la materia.	V	F	NS
4	Las gotitas de agua y/o hielo que contiene el aire forman las nubes.	V	F	NS
5	La densidad es la relación entre la masa y el volumen.	V	F	NS
6	Los gases se componen de partículas que se mueven libremente.	V	F	NS
7	El movimiento de traslación de la Tierra da lugar al día y la noche.	V	F	NS
8	Cuando un gas se comprime el tamaño de las partículas que lo componen también disminuye.	V	F	NS
9	La Tierra está en el centro del Universo.	V	F	NS
10	Todo gas cerrado en un recipiente provoca una fuerza sobre sus paredes.	V	F	NS
11	1 litro de volumen de cualquier líquido es igual a 1 kg de peso.	V	F	NS
12	Una partícula de un gas pesa siempre igual independientemente de su temperatura.	V	F	NS
13	La unidad mml/cm mide la densidad.	V	F	NS
14	Las partículas de los gases están en constante movimiento.	V	F	NS
15	La presión del aire se ejerce sólo hacia abajo.	V	F	NS
16	La atmósfera está compuesta por partículas en estado gaseoso.	V	F	NS
17	La Tierra es el planeta más cercano al Sol.	V	F	NS
18	Las partículas de los gases se mueven más deprisa por el efecto del calor.	V	F	NS
19	Un litro de mercurio pesa lo mismo que un litro de agua.	V	F	NS
20	Cuando disminuye la temperatura de un gas disminuye su presión.	V	F	NS
21	El peso es la fuerza con la que la Tierra atrae un cuerpo.	V	F	NS
22	El peso de un gas dentro de un recipiente depende de la temperatura a la que se encuentre.	V	F	NS
23	Las mareas se deben principalmente a la atracción que la Luna y el Sol ejercen sobre la Tierra.	V	F	NS
24	Un gas sólo puede estar formado por partículas idénticas.	V	F	NS
25	Cuando ponemos en contacto un cuerpo frío y otro caliente, la energía pasa del cuerpo caliente al frío.	V	F	NS
26	Al comprimir un gas aumenta su densidad.	V	F	NS
27	Podemos conocer el volumen de un sólido si lo sumergimos en un líquido y medimos la cantidad de líquido que se desplaza.	V	F	NS
28	El tipo de partículas que componen un gas no influye en su peso.	V	F	NS
29	Los cuerpos en equilibrio térmico tienen la misma temperatura.	V	F	NS
30	El hielo flota en el agua porque su densidad es mayor.	V	F	NS

TEXTO FASE APRENDIZAJE.

LA PRESIÓN ATMOSFÉRICA Y EL VIENTO

La Tierra está rodeada por una capa de gases que la separa del espacio vacío que constituye, en su mayor parte, el Universo. Esta capa recibe el nombre de atmósfera y está formada por una mezcla de gases que llamamos aire.

En esta unidad aprenderás algunas características básicas de la atmósfera, entre ellas, qué es presión atmosférica y uno de los fenómenos más importantes relacionados con esta: el viento.

La presión atmosférica

La atmósfera está formada por un conjunto de gases que, como toda materia, está compuesto por partículas que tienen masa y, por lo tanto, peso. Se estima que el aire que compone la atmósfera pesa, aproximadamente, 5.500 billones de toneladas.

La presión atmosférica es la fuerza que se ejerce, en un punto concreto, por el peso de la columna de aire que se extiende por encima de ese punto, hasta el límite superior de la atmósfera.

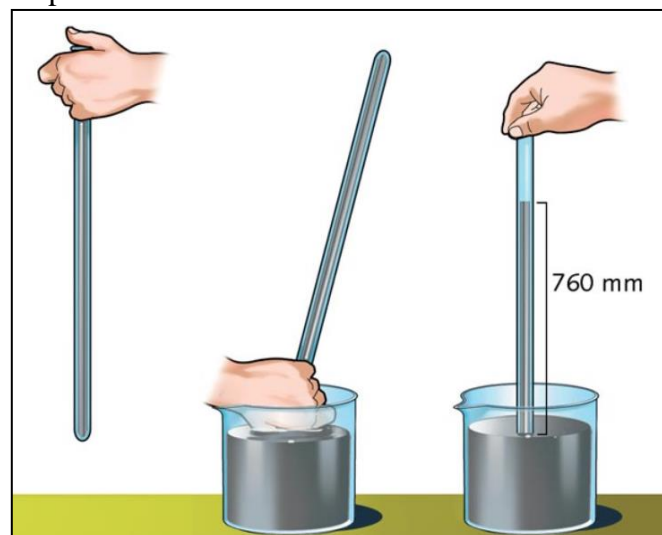
Todos los materiales y los seres vivos que poblamos este planeta estamos sometidos a la presión atmosférica. Si no somos conscientes de este peso del aire es porque ya estamos adaptados porque se ejerce por igual en todas direcciones y nuestros líquidos internos están a la misma presión.

El descubrimiento de la presión atmosférica.

La fuerza del peso del aire no se descubre hasta 1643. Fue el científico italiano Evangelista Torricelli quien ese año hizo un experimento que demostró que el aire ejercía presión.

La imagen 1 muestra cómo lo hizo: en primer lugar, llenó con mercurio un tubo de un metro de longitud cerrado por un extremo y tapó el extremo abierto con el dedo. Luego, lo introdujo invertido en una cubeta llena de mercurio y, finalmente, retiró el dedo con cuidado para que no entrara aire en el tubo. En ese momento, el mercurio descendió hasta una altura de 760 mm sin llegar a vaciar el tubo y dejó un vacío en el extremo cerrado del tubo. Así fue como Torricelli también fue el primero en establecer cómo medir la presión atmosférica.

Imagen 1. Experimento de Torricelli



En la antigüedad, se creía que el vacío no era algo natural y que la naturaleza se resistía a tolerar la ausencia de aire. Torricelli demostró con su experimento que no es esa resistencia al vacío lo que impide al mercurio salir del tubo, sino que es la atmósfera la que ejerce presión sobre el mercurio de la cubeta. Por eso, la columna de mercurio desciende hasta una altura de 760 mm, el punto que iguala la presión que el aire ejerce sobre el mercurio de la cubeta.

El aparato que diseñó Torricelli es el barómetro y se utiliza para medir la presión atmosférica. La presión del aire a nivel de mar equivale a la presión que ejerce una columna de mercurio (Hg) de 760 mm de altura, valor que equivale a una atmósfera (760mmHg = 1 atm). En meteorología se suelen utilizar otras unidades de medida como son el hectopascal (hPa) y el milibar (mb).

La presión atmosférica varía

Como sabes, en los gases las partículas se mueven libremente ocupando todo el volumen. Pero en la atmósfera, al ser un espacio tan grande, la distribución de las partículas que conforman el aire no es uniforme, debido a que las condiciones de cada lugar son diferentes. Así, la densidad de las capas de aire varía según la altura, ya que el tamaño de la capa de aire por encima es diferente. Por eso, en las capas inferiores de la atmósfera, que soportan el peso de todas las que están encima, las partículas de aire se comprimen más, se mueven menos y el aire es más denso (hay más cantidad de partículas de gas por unidad de volumen).

Como la presión depende del peso del aire que tenemos por encima, a medida que ascendemos la presión va disminuyendo. Por eso, la presión que existe en la cima de una montaña de 3.000 m de altitud es menor que la de una playa. Como la presión atmosférica varía con la altura se ha establecido que la presión normal, la equivalente a 1 atm (1013 hPa), es la que se da justo al nivel del mar, porque es la referencia que también utilizamos para indicar la altura de cualquier punto geográfico. Las presiones superiores a ésta se denominan altas presiones y las inferiores, bajas presiones.

La presión atmosférica además de variar con la altura también varía con la temperatura. Cuando los gases se calientan sus partículas se aceleran y tienden a expandirse, separándose y reduciendo su densidad, lo que afectaría a su presión; lógicamente ocurre todo lo contrario cuando se enfrían.

La presión atmosférica y el viento

Las diferencias en la densidad y presión atmosférica de las distintas partes del planeta son las responsables de los vientos y de otros fenómenos meteorológicos.

Las zonas de baja presión atmosférica, llamadas de ciclón o borrasca (imagen 2), se forman por masas de aire caliente que cuando ascienden dejan tras de sí un área de baja densidad. Mientras que en las zonas de anticiclón (imagen 3) son las masas de aire frío, más denso, las que tienden a descender desde las capas altas; causando la compresión de las masas de aire inferior, dando lugar a zonas de alta presión atmosférica.

Imagen 2. Formación de una zona de baja presión o borrasca. El aire caliente asciende y el hueco que deja lo llenan masas de aire vecinas.

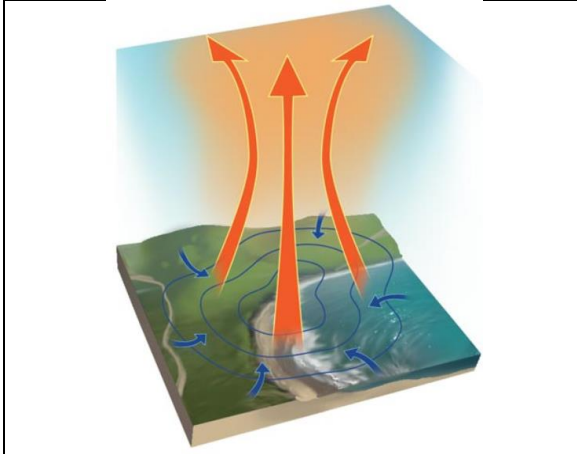
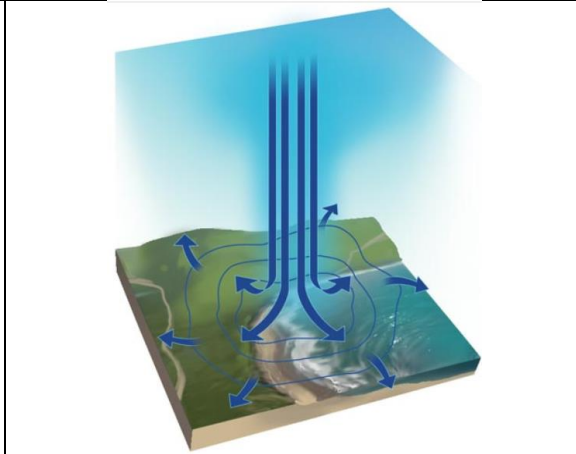


Imagen 3. Formación de una zona de alta presión o anticiclón. El aire frío desciende y comprime el aire inferior que se dispersa al llegar a la superficie.



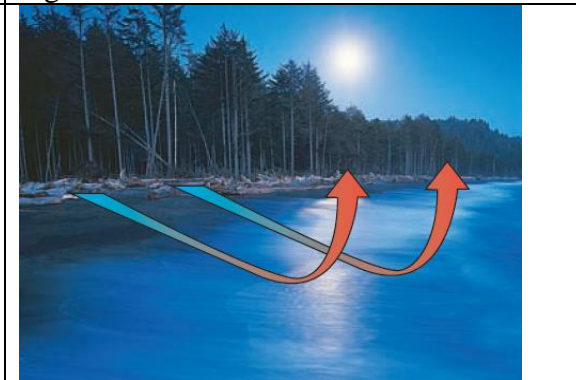
Si combinamos estos dos fenómenos podemos entender cómo funciona la dinámica de la atmósfera y cómo se produce el viento. El viento es el movimiento de grandes masas de aire a través de la troposfera (la capa inferior de la atmósfera). La existencia de zonas de baja y alta presión ocasiona movimientos de aire que tienden a igualar las presiones de las distintas áreas, provocando corrientes de aire que van desde las zonas de alta presión hasta las zonas de baja presión.

De la misma manera se producen otros fenómenos que son más locales y tenues como por ejemplo, las brisas diurnas y las nocturnas (imágenes 4 y 5).

Imagen 4. Durante el día, la tierra se calienta más deprisa que el mar. El aire caliente de la costa asciende y es sustituido por el aire más frío procedente del mar, proceso que da lugar a la brisa diurna.

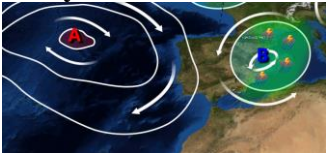


Imagen 5. Durante la noche, la tierra se enfría más deprisa que el mar. El aire sobre el mar, más caliente, asciende y el lugar de éste es ocupado por el aire más frío procedente de la costa, hecho que da lugar a la brisa nocturna.



PREGUNTAS FASE APRENDIZAJE Y MENSAJES EF

LA PRESIÓN ATMOSFÉRICA Y EL VIENTO	
Preguntas	EF (Explicación)
<p>1. ¿Qué da lugar a la presión atmosférica?</p> <p>a) <u>El peso de los gases que contiene la atmósfera.</u> b) La expansión de los gases a través de la atmósfera. c) El aumento del volumen de las partículas de aire. d) La resistencia a tolerar la ausencia de aire.</p>	<p>Piensa que la atmósfera está compuesta de materia y ésta la debes ver como una sucesión de capas de aire, unas sobre otras aplastándose.</p>
<p>2. ¿Por qué en el experimento de Torricelli el mercurio sólo desciende hasta los 760 mm?</p> <p>a) Porque se crea un vacío en la parte superior del tubo que impide que siga cayendo. b) Porque la presión de la atmósfera no es lo suficientemente grande para empujar el mercurio más abajo. c) <u>Porque a esa altura la fuerza que ejerce la masa del mercurio en la barra es igual a la fuerza que ejerce la columna de aire sobre la cubeta.</u> d) 240 mm de mercurio es la cantidad que necesitamos para igualar la densidad del mercurio de la cubeta a la densidad del aire.</p>	<p>Piensa que el mercurio sale del tubo mientras su peso es mayor que el peso del aire exterior que hay por encima. Y deja de salir cuando el peso del mercurio del tubo y el peso del aire se igualan.</p>
<p>3. Si Torricelli hubiera cogido un tubo de 2 metros de alto lleno de mercurio, en vez de un tubo de un metro, y hubiera repetido su experimento, ¿qué habría pasado con la columna de mercurio?</p> <p>a) Habría quedado a 1760 mm de altura. b) Habría quedado al doble de altura, a 1520 mm. c) Habría quedado más abajo de 760 mm. d) <u>Habría quedado a la misma altura, 760 mm.</u></p>	<p>Piensa que la altura a la que queda el mercurio está determinada por el punto en que se iguala el peso que ejerce la columna de aire sobre la cubeta y el peso del mercurio que contiene el tubo.</p>
<p>4. Si intentas replicar el experimento de Torricelli en lo alto de una montaña de ocho mil metros, en vez de al nivel del mar, donde se realizó el original ¿Qué crees que sucederá?</p> <p>a) <u>Que saldrá más mercurio del tubo a la cubeta, bajando más de 760 mm de altura.</u> b) Que saldrá menos mercurio desde el tubo a la cubeta, bajando menos de 760 mm de altura. c) Que saldrá la misma cantidad de mercurio del tubo a la cubeta, quedando a 760 mm de altura. d) Que debido a la altura saldrá casi completamente el mercurio del tubo.</p>	<p>Piensa que como estás a más altura hay menos cantidad de aire por encima; por eso, el peso del aire y del mercurio del interior del tubo se igualará cuando haya salido una cantidad diferente de mercurio que cuando se hace a nivel de mar.</p>
<p>5. ¿La densidad del aire es igual en todos los puntos de la tierra?</p> <p>a) Sí, porque los gases se expanden y tienden a ocupar todo el espacio. b) <u>No, porque varía en función de la temperatura y la altura.</u> c) No, porque el viento hace que varíe la densidad del aire. d) Sí, porque la mezcla de gases que componen la atmósfera es siempre igual.</p>	<p>Piensa que la densidad del aire se determina por lo juntas o separadas que se encuentran las partículas que lo componen. Y eso depende de las condiciones externas.</p>

<p>6. ¿Por qué en la atmósfera pesa más el equivalente al volumen de un litro de aire frío que el de un litro de aire caliente?</p> <p>a) Porque las partículas frías pesan más. b) <u>Porque hay más partículas al estar más juntas.</u> c) Porque el aire caliente disminuye su volumen. d) Porque las partículas del aire se dispersan al enfriarse.</p>	<p>Piensa que la temperatura influye sobre la velocidad de movimiento de las partículas de un gas y, por tanto, influye en lo más o menos próximas que pueden estar. Así, dos volúmenes iguales de aire a distinta temperatura tendrán densidades diferentes.</p>
<p>7. Imagina que llenamos una botella de plástico flexible con aire bastante caliente y la cerramos herméticamente ¿Qué pasará con la botella cuando se enfríe?</p> <p>a) La botella se hinchará y aumentará su tamaño al enfriarse el aire del interior. b) La botella pesará más al enfriarse el aire del interior, el cual se hará más denso y pesado. c) La botella pesará menos, ya que al enfriarse el aire se hará menos denso y pesado. d) <u>Que la presión atmosférica chafará la botella ya que al enfriarse el aire del interior reducirá su presión.</u></p>	<p>Piensa que al enfriarse el aire cambia la relación de presión entre el interior y el exterior, ya que las partículas pasan a moverse más lentamente en el interior de la botella.</p>
<p>8. Imagina un globo aerostático abierto por la base, completamente hinchado y ascendiendo al calentar el aire del interior con un quemador. ¿Qué ocurre con la densidad del aire dentro del globo?</p> <p>a) Es igual a la del aire exterior porque el globo está abierto. b) Es mayor porque aumenta el tamaño de las partículas. c) <u>Es menor porque el aire está más caliente.</u> d) Es menor porque el volumen del globo disminuye.</p>	<p>Piensa que cuando calentamos el aire interior hacemos que sus partículas se muevan con más energía lo que las obliga a expandirse y a escapar del globo al estar abierto por la base.</p>
<p>9. ¿Por qué varía la presión atmosférica con la altura?</p> <p>a) Porque a nivel del mar la presión se ve afectada por la brisa diurna y nocturna. b) <u>Porque cambia el peso de la columna de aire que tiene por encima.</u> c) Porque en la cima de una montaña de 3000 metros la presión es mayor al hacer más frío. d) Porque la presión atmosférica normal es la que se da a nivel del mar.</p>	<p>Piensa en la atmósfera como una serie de capas de aire que se superponen unas sobre otras, y el número de capas que hay encima varía según la altura a la que nos encontramos.</p>
<p>10. ¿Por qué se origina una borrasca?</p> <p>a) Porque el aire de la superficie es más denso que el aire que le rodea, haciendo que descienda. b) Porque el aire de la superficie es tan denso como el que le rodea pero está más frío. c) <u>Porque el aire de la superficie es menos denso que el aire que le rodea, haciendo que ascienda.</u> d) Porque el aire de la superficie es tan denso como el que le rodea pero está más caliente.</p>	<p>Piensa que en el centro de una borrasca existe una masa de aire en superficie con menor cantidad de partículas que las masas de aire que le rodean. Y, por tanto, la masa de aire central está rodeada de aire con mayor cantidad de partículas que le presionan.</p>
<p>11. ¿Qué ocurriría en España si hubiera una borrasca en el Mediterráneo y un anticiclón en Portugal?</p> 	<p>Piensa que el viento se mueve siempre desde zonas de alta presión a zonas de baja presión. Como en las zonas de baja presión el aire es menos denso, asciende, tendiendo a ser ocupadas por aire más denso procedente de zonas de</p>

<p>a) En España al estar fuera de la zona de borrasca y de anticiclón habría un tiempo estable y sin viento.</p> <p>b) En una distancia tan amplia no influiría una sobre la otra.</p> <p>c) Habría movimientos de masas de aire para que la borrasca y el anticiclón se fueran separando.</p> <p>d) <u>Habría movimientos de masas de aire desplazándose desde Portugal hacia el Mediterráneo.</u></p>	<p>alta presión.</p>
<p>12. En la orilla del mar, durante el verano a mediodía, a veces podemos notar un poco de brisa ¿A qué se debe?</p> <p>a) <u>A la entrada de aire que proviene del mar refrescando la playa.</u></p> <p>b) A la salida de aire del interior hacia el mar refrescando la playa.</p> <p>c) A los movimientos de grandes masas de aire debidos a la presión.</p> <p>d) A que la tierra después del mediodía está más fría que el mar.</p>	<p>Piensa que durante el día, el aire del interior se hace menos denso al variar su temperatura por lo que tenderá a ascender y su espacio se ocupará por aire más denso.</p>

PREGUNTAS POST-TEST (Fase final)

1. ¿A qué conclusión fundamental llegó Torricelli tras su experimento en relación a la presión atmosférica?
2. ¿Qué sucede con las partículas de un gas cuando se calienta?
3. Imagina dos globos elásticos idénticos llenos de la misma cantidad de aire y cerrados herméticamente. Si subimos 10° la temperatura de uno de ellos (A), pero bajamos 10° la del otro (B) ¿variara el tamaño de ambos? ¿Por qué?
4. ¿Qué es el barómetro?
5. ¿Siempre que haya zonas con diferente presión atmosférica habrá viento? ¿Por qué?
6. ¿En qué condiciones habrá más olas en dirección a la playa, (A) cuando hay bajas presiones en la costa y altas mar adentro, o (B) cuando hay altas presiones en la costa y bajas mar adentro? ¿Por qué?
7. ¿Dónde es mayor la presión atmosférica, en lo alto de una montaña de 2000 metros o en la orilla de la playa?
8. Imagina que Alicia está en la orilla del mar y María en un barco a 1 km de la costa y son las 2 h. de la tarde. ¿Cuál de las dos amigas notará una brisa agradable? ¿Por qué?
9. ¿Cómo se puede hacer que un globo aerostático ascienda o descienda de altura? ¿Por qué?
10. ¿Qué factores hacen variar la presión atmosférica?

ANEXO C

TEXTO Y PREGUNTAS DEL ESTUDIO 4

El cuestionario de conocimientos previos, así como el texto y las preguntas administradas en la fase de aprendizaje, son idénticos a los empleados en el estudio 3 (Ver ANEXO B).

PREGUNTAS POST-TEST (Fase final)

1. Imagina que Torricelli hubiera dispuesto de oro líquido, el cual es bastante más denso que el mercurio, y lo hubiera introducido en el tubo. ¿Qué crees que habría pasado?
 - a) El oro habría quedado a la misma altura, 760 mm.
 - b) El oro habría quedado más abajo de los 760 mm.
 - c) Se habría salido prácticamente todo el oro.
 - d) No habría salido casi oro quedando cerca de los 1000 mm.

2. ¿La fuerza que ejerce la atmósfera es igual en todos los puntos de la tierra?
 - a) Sí, porque la composición de las partículas del aire es uniforme en todos los puntos.
 - b) Sí, porque esta fuerza se ejerce por igual en todas las direcciones.
 - c) No, porque depende del peso del aire de las capas inferiores.
 - d) No, porque la densidad del aire depende de la altura de ese punto.

3. ¿Cómo varía la densidad del aire de la atmósfera en lo alto de una montaña?
 - a) Es menor porque la columna de aire que tenemos por encima es menor.
 - b) Es mayor porque al estar más altos estamos más cerca de las capas de aire.
 - c) Es mayor porque el aire tiene menos humedad.
 - d) Es menor porque la temperatura es más baja.

4. Si quisiéramos subir una gran montaña, como el Everest, veríamos que nos cuesta mucho respirar. ¿Por qué crees que sucede esto?
 - a) Porque el aire al estar muy frío es muy pesado y cuesta respirarlo.
 - b) Porque la presión del aire es mayor e impide que podamos respirar bien.
 - c) Porque hay menos aire, ya que este es menos denso a esa altura.
 - d) Porque a esa altura el aire es más denso y empieza condensarse.

5. ¿Qué hace que el aire caliente ascienda?
 - a) Que haya una zona de baja presión atmosférica.
 - b) Que tiende a expandirse para ocupar todo el espacio.
 - c) Que tiene una densidad menor que el aire frío.
 - d) Que las partículas de aire se vuelvan más ligeras.

6. ¿Qué pasará con un globo elástico hinchado y cerrado herméticamente si sube la temperatura y se calienta el aire de su interior?

- a) Al calentarse aumentará su volumen.
- b) Al calentarse disminuirá su volumen.
- c) Al calentarse pesará menos y ascenderá.
- d) Nada, ya que el número de partículas dentro del globo permanece constante.

7. ¿Cómo se explica el viento?

- a) Por los movimientos del aire que tienden a enfriar la temperatura del planeta.
- b) Por el desplazamiento del aire desde zonas de alta a las de baja presión.
- c) Por el movimiento de rotación de la troposfera alrededor de la tierra.
- d) Por el proceso de mezcla de los gases que componen la atmósfera.

8. En las zonas costeras, por la noche, podemos notar que la temperatura no es tan baja como en las zonas de interior ¿a qué se debe este fenómeno?

- a) A que la humedad existente en el ambiente en las zonas costeras produce que la sensación térmica sea más cálida y existan menos variaciones de temperatura.
- b) A que la presión atmosférica en las zonas de costa es más estable que en las zonas de interior por lo que existen menos movimientos de masas de aire que las enfríen.
- c) A que en el interior hay una mayor altitud por lo que son alcanzadas antes por las masas de aire frío que bajan desde las capas altas de la atmósfera.
- d) A que el mar se enfría más despacio que la tierra; así, por la noche el aire caliente sobre el mar asciende dirigiéndose hacia el interior de la costa, mientras que el aire más frío de la costa se dirige por la capa inferior hacia el mar calentándose.

El autor de la presente tesis ha disfrutado de una beca de Formación de Profesorado Universitario (FPU) (FPU14/04646) del Ministerio de Educación, Cultura y Deporte. Asimismo, ha disfrutado de dos becas para realizar estancias breves en la University of Minnesota (EST15/00492) y en la Université de Poitiers (EST16/00323), financiadas por el Ministerio de Educación, Cultura y Deporte.