VNIVERSITAT DE VALÈNCIA | LECTURA Estructura de Recerca Interdisciplinar

Faculty of Psychology

Doctoral Programme in Reading and Comprehension

WITHIN THE IBM GALAXY: EXPLORING DIFFERENCES IN ONLINE PROCESSES AND COMPREHENSION OUTCOMES BETWEEN READING ON SCREEN AND READING ON PAPER

DOCTORAL THESIS WITH INTERNATIONAL DOCTORATE MENTION

Presented by
PABLO DELGADO HERRERA

Supervised by
LADISLAO SALMERÓN GONZÁLEZ, PhD

Valencia, January 2020

VNIVERSITAT DE VALÈNCIA

FACULTY OF PSYCHOLOGY

Doctoral Programme in Reading and Comprehension

WITHIN THE IBM GALAXY:
EXPLORING DIFFERENCES IN ONLINE PROCESSES AND COMPREHENSION OUTCOMES BETWEEN READING ON SCREEN AND READING ON PAPER

DOCTORAL THESIS WITH INTERNATIONAL DOCTORATE MENTION

Presented by

PABLO DELGADO HERRERA

Supervised by

LADISLAO SALMERÓN GONZÁLEZ, PhD

Valencia, January 2020

This doctoral dissertation was funded by a grant (BES-2015-076055) from the Spanish Ministry of Science, Innovation, and Universities and the European Social Fund, as part of the project *Lectura Crítica en Internet: Evaluación, Desarrollo e Intervención* (EDU2014-59422-P)

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

European Union
European Social Fund

A mamá, a papá, a Fran y a Gloria.

[This page intentionally left blank]

Acknowledgements - Agradecimientos

A mi director, Lalo. Muchas gracias por enseñarme este oficio. Por ser mi guía y hasta mi asistente. Por enseñarme que una premisa básica es estar dispuesto a ser convencido y por ayudarme a ver más allá del alcance de mis ojos. Por todo el esfuerzo que me has dedicado. Por implicarme. Por tu tolerancia con mis errores. Gracias.

A Cristina, no solo por tu imprescindible trabajo en la realización del meta-análisis, sino por tu ayuda durante estos cuatro años y, en especial, por tu cariño. Gracias.

A Inma y a Vicen, por hacer siempre que el trabajo junto a vosotras fuera tan fácil. Por vuestro apoyo. Por contar conmigo. Gracias.

Gracias, Javier, por estar siempre dispuesto a ayudarme y por tu exquisita amabilidad.

A David Saldaña, por confiar en mí cuando comencé mi andadura académica. Por tu ayuda y guía durante mis primeros pasos. Gracias.

A mis compañeros Arantxa y Nacho, por vuestro apoyo y compañía. Ha sido un placer, de verdad. Gracias, Arantxa. Gracias, Nacho.

Gracias, Marian, por ser tan cariñosa. Por tu ayuda. Por nuestras charlas y risas. Ahora echo de menos que hubieras estado todos estos años más cerca.

Ivar, Helge, Øistein, and Elisabeth, I am particularly grateful for your hospitality during my stay in Oslo. Ivar, I would also like to thank you for your work and your insights. Helge, thank you again, for your help and warmth.

Rakefet, my grateful thanks are also extended to you, for your work and teachings.

A Laura Gil, Amelia y Antonio Ferrer, por vuestros consejos y por acogerme con tanto cariño desde el primer día. Gracias, Antonio, por ayudarme a salir de aquel embrollo.

Muchas gracias, Nadina, Marina, Bea, y Alba por vuestra ayuda y por algunos ratitos de risas. También los he tenido con vosotras, Laura Royo y Vittoria, quienes además me habéis ayudado con la recogida de datos. Gracias.

A Raquel, por tus gestiones administrativas durante estos últimos meses.

A Manolo, también por alguna que otra ayuda.

Gracias, Bea Lucas, por tu sonrisa por los pasillos.

A Rafa y a Paco, por vuestra dirección del Departamento. Y a Sergio, Toni, Miguel y Pepe por vuestro trabajo y ayuda desde la Secretaría.

Gracias a la editorial Prensa Científica, por cederme desinteresadamente 10 copias de su revista Investigación y Ciencia. Gracias también a la Fundación Infancia y Aprendizaje y a la traductora Mary Black, por cederme la traducción del artículo incluido en el Capítulo 2.

I would also like to thank Faye Antoniou, Mirit Barzillai, Gal Ben-Yehudah, Susana Padeliadu, and Kate Ziegelstein for their contribution to the meta-analysis. And Jukka Hyönä, who spent his time reviewing one of the studies that take part of this thesis.

Además, no me gustaría olvidarme de quienes me hacen feliz alejados de la Universidad:

A Billy, Kroata y Cromagnonator. Habéis conseguido que mi vida en Valencia sea mucho más divertida. Me habéis acompañado mucho. En realidad, los tres sois mis favoritos.

Gracias, Nico. Hace mucho que no te veo, amigo.

Gracias, Ignacio. Te echo mucho de menos. Espero que la vida nos junte de nuevo, y que seas feliz.

A Jorge, José Ramón y Valle por hacerme sentir querido. Gracias.

Gracias, Fran, amigo mío. Mi hermano. Qué lejos estamos, pero cuánto nos queremos. El rubito ese, ya lo sabes, va a hacer carrera.

Gracias, Gloria, por tu amor. Por haberme impulsado desde el primer instante, a pesar de que debía marcharme a más de 600 kilómetros. Por tu apoyo. Por ser mi compañera. Eres bonita haciendo cualquier cosa. Te amo.

Gracias, papá, por tu amor y apoyo sin condiciones. No estaría aquí sin ti. Te amo.

Mamá, gracias por tu dedicación infinita. Por tu amor sin condiciones. No estaría aquí sin ti. Te amo.

Por último, gracias a todas las personas que aportáis al bien común en España. Me siento un privilegiado cada vez que pienso que habéis pagado mi trabajo con el vuestro. Me impulsáis a esforzarme para que haya merecido la pena.

Gracias

“*Perhaps this is the way to help those who want to say: “But is there nothing good about print?” The thing of this book is not that there is anything good or bad about print but that unconsciousness of the effect of any force is a disaster, especially a force that we have made ourselves.”*

Marshall McLuhan, *The Gutenberg Galaxy* (1962)

[This page intentionally left blank]

Epigraph

IN 1962, MARSHALL McLuhan published The Gutenberg Galaxy, one of the most popular books in communication research. Virtually ignoring the content of the message, the author reviewed the decisive influence of the massive practice of printed reading on the human being, from sociological to philosophical, cultural, and even psychological consequences. McLuhan analysed the *typographic* man and its society, compared to the oral man from the humanity prior to the widespread use of the Gutenberg moveable type printing press. His work did not obey a random whim, but was a corollary of the advent of the electronic mass media, which, according to him, were partially returning the western society to the ancient orality. McLuhan coined the popular term *global village* to refer to the society interconnected by mass media, and the well-known aphorism *the medium is the message.* His ideas, albeit not exempt from extravagances and well-founded criticisms, seem to be particularly relevant in today's Digital Era.

The research domain and the methodology of the present dissertation are distant from McLuhan's realm, but his comments were an inspiration. The present dissertation is based on the idea that the way we usually interact with communication technology shapes the manner in which we approach and process the information. Specifically, the following chapters shall address the influence of the reading medium (print vs. screen) on reading performance. McLuhan, who was died in 1980, could barely see the ineffable changes, both at collective and at individual level, caused by the Digital Era over the last four decades. His *typographic* man, who represented the consequences of a complex network of psychological and social effects that the author called the *Gutenberg Galaxy*, is now mutating into what we could call the *pixelgraphic man*, this is, the sons and daughters of the *IBM Galaxy*. The consequences of this new era are overwhelmingly vast and complex. We shall explore a small portion of the picture.

Index

[This page intentionally left blank]

Summary

THE PRESENT DISSERTATION aimed to explore the influence of the reading medium on reading performance. It consists of six chapters, including a general introduction, a brief narrative review, three empirical studies, and a general discussion. In Chapter 1 (*General Introduction*) we introduce the topic. As a point of departure, we briefly discuss the existing controversy about the use of digital technologies in education. We subsequently focus on the influence of the digital medium on text comprehension. To this end, we first review some of the most influential models of reading comprehension. Then, we review previous findings showing that people still prefer printed texts, especially when it comes to in-depth reading. In this regard, we shall see that people's impressions usually point out that reading on screens disrupts attention and concentration. We then argue that the type of interactions that people usually have with digital technologies may be responsible for the detrimental effect of the digital medium previously found by some studies. In this regard, we present the *Shallowing hypothesis* of the medium effect, which proposes that reading on screen hinders reading comprehension by fostering cognitive disengagement, so that the processing of the information becomes superficial.

In Chapter 2 we include a narrative review that goes further into findings showing that digital technologies are not always suitable for learning. Subsequently, we present in Chapter 3 the first empirical study of this dissertation, which consists of a meta-analysis including the existing literature from 2000 to 2017 that compares reading comprehension across reading media. The results showed poorer comprehension outcomes when reading on screen. Furthermore, some relevant moderators qualified this effect. The on-screen

inferiority was found to be larger among those studies that used expository texts (vs. narrative texts) and those that required participants to read under time constraints (vs. self-paced reading time). These findings pointed out that the medium effect especially appears when the reading task demands increased mental effort. In addition, we found that the effect is larger among the more recent studies, so in younger generations.

Based on the results above, we conducted two experimental studies aiming to test the Shallowing hypothesis of the medium effect. Chapter 4 includes a study that compared readers' engagement between reading media through their eye movements. It also examined readers' comprehension and meta-comprehension of several expository texts. A sample of 116 undergraduates read tree texts on a printed booklet and three texts on a tablet. Some of the participants self-paced their study time, whereas the rest of the sample read under time pressure. The findings indicated that, regardless of the reading-time frame, participants fixated longer when reading in printed texts, and that they were more accurate in this medium when monitoring their comprehension. Accordingly, they scored higher on the comprehension questions, although this effect only approached significance with a conservative two-tail test.

Our third study (Chapter 5) compared readers' on-task attention, comprehension and meta-comprehension when reading a long expository article in print vs. on screen. One hundred and forty undergraduates were allocated to one of four experimental condition, varying in the reading medium (desktop computers vs. the actual printed magazine) and in the reading time-frame (self-paced vs. time pressure). The results showed that only those participants who read in print reduced their mindwandering when the task required to read under time pressure. Thus, in this time-frame condition, the participants who read on screen scored lower on the comprehension test. In contrast, when reading time was self-paced, the comprehension scores were similar regardless of the medium. Finally, there were no differences in metacognitive monitoring of comprehension between the experimental groups.

Finally, in Chapter 6 we discuss the overall findings from these three empirical studies. We conclude that they support the Shallowing hypothesis. We also suggest future research lines and we discuss the main implications of our findings.

Chapter 1

General Introduction

~

"Books will soon be obsolete in the public schools.”

Thomas A. Edison (1912)

"These data show that the reality in schools lags considerably behind the promise of technology.”

Andreas Schleicher (2015)

Chapter 1

General Introduction

THE INCLUSION OF communication technologies as instructional tools has been constant since the first decades of the 20th century. From the audio-visual instruction movement during the 1920's and the 30's to the use of computers, tablets, and the Internet, practitioners and policy makers have been generally enthusiastic about the electronic technologies in education (for an overview see Molenda, 2008). However, their learning outcomes have been controversial from the very beginning. In this chapter, we will first briefly look at the relationship between digital technologies and formal education. Then, we shall focus on the specific influence of the use of digital devices for reading. To that end, we shall first review some of the most influential theoretical models of reading comprehension. Then, we shall see that, in spite of the pervasive use of digital devices for reading, people still prefer printed texts when they aim to read in depth. They often indicate that reading on screens hinders attention and in-depth reading. We will argue that our usual interactions with the digital technologies could be shaping the way we read on them, so that they may not be as suitable as printed texts for reading comprehension. In this direction, the *Shallowing hypothesis* proposes that reading on screen fosters a superficial processing of the information.

1.1. A brief overview of the use of digital media in education

The debate on the usefulness of electronic media in education especially increased during the 80s and the 90s. In one of the most relevant reviews on this issue, Clark (1983) drew attention to the absence of media effects on learning above and beyond the instructional methods. Although several studies reported fruitful learning outcomes, Clark insisted on that they were not an exclusive contribution of the electronic media, but of the pedagogical approach or the effect of novelty on students' engagement. Thus, according to his critical analysis, the author strongly concluded that "media do not influence learning under any conditions" (Clark, 1983, p. 445; see also Clark, 1994). Contrary to Clark's view, Kozma (1991) defended that the unique capabilities of media, together with the instructional methods specially designed for them, could result in different processes influencing learning outcomes. In an extensive review of primary studies, the author tried to support that learning through media –i.e., books, television, computers, and multimedia (as an extension of computers)– yields differentiated cognitive effects. He pointed out that medium influences, either in a positive or a negative way, the creation of learners' mental models of the learning content. According to this author, it is wrong to separate medium from method because they both are part of the same thing: the design. Media interacts with methods (and vice versa) so that learners relate to them in a specific manner. Kozma also claimed that some methods are not possible to be implemented in any medium, particularly those allowed by the attributes of computers. Not surprisingly, the author called for extending research on the use of computers and, especially, on the implementation of the hypertext and hypermedia, the then emergent way of text presentation in the field of instructional designs and professional environments (Kozma, 1991).

According to Kozma (1991, 1994), the burgeon of the digital technology and the Internet took a step further into the possible effects of electronic media on learning. During those first years of use of digital technologies, scholars expressed strong convictions about the fact that they represented an unprecedented opportunity for instruction and learning, especially when it comes to learning complex content (e.g., Lanham, 1995; Spiro & Jehng, 1990). Nevertheless, some concerns about their possible negative effects also began to grow based on the idea that the digital environments were hindering people's learning, especially through their reading experiences. Those initial warnings suggested that reading on screens would prevent learners to engage with texts (e.g., Birkerts, 1994; Healy, 1999).

Since then, Western society has evolved dramatically in terms of the presence and the use of digital technologies. Today's pervasive use of mobile devices keeps anyone connected to anything, anywhere, anytime. In the realm of education, the presence of computers in schools has not stopped increasing. Schools in the USA had one computer for every 14 students in 1992 (Anderson & Ronnkvist, 1999), and only 35% had Internet access in 1994 (Bare & Meek, 1998). Nevertheless, as early as in 1998, computer-student proportion in USA increased to 1/6, 90% of schools had Internet access (Anderson & Ronnkvist, 1999), and half of the teachers used them regularly in their instructional practices (Smerdon et al., 2000). In 2012, almost 95% of high school students in OECD countries had at least one computer at home, and almost 90% had Internet connection. With regard to their use for formal learning, the average computer-student rate was 1/4 in the OECD countries (1/2 in the USA), slightly more than 70% of students used computers regularly at school, and 64% connected to Internet during a typical school day (OECD, 2015).

However, in spite of this widespread presence of computers and other digital devices in education, their benefits for learning are still controversial. On the one hand, some meta-analyses have found small to moderate learning benefits from the use of technology-based interventions compared to control groups. For example, Schmid et al. (2009) found an effect size (ES) of 0.28 in a sample of 231 studies conducted in higher education. A second-order meta-analysis by Tamim, Bernard, Borokhovski, Abrami, and Schmid (2011), which included meta-analyses on different methodologies, subjects, and populations (from K-12 to college students), reported an ES of 0.33 also favoring technology-based instruction. Both meta-analyses concluded on the role of the pedagogical approach, because the technology-based interventions aiming cognitive support (e.g., planning, monitoring, meta-reasoning) yielded better results than those used to present or deliver subject content. Nonetheless, the authors concluded that the high heterogeneity regarding methodologies, types of control groups, subjects, and populations made hard to generalize the results (Schmid et al., 2009; Tamim et al. 2011).

Two meta-analyses by Cheung and Slavin (2012, 2013) focused on reading and on mathematics interventions, respectively, also found small effects favouring technology-based instruction (ES = 0.16 and 0.15). Of note is that the interventions that yielded better results in the meta-analysis on reading were those that mixed non-technology and computer-assisted instruction. Furthermore, results showed that the higher the methodological quality of the studies the lower the effect (Cheung, 2012). Recently, Haßler, Major, and Hennessy (2016) conducted a narrative review on the use of tablets in educational practices at schools, reporting that the majority of the studies (16 out of 23) yielded positive learning outcomes. The authors were also cautious about the generalizability of the results, due to the low methodological quality of the studies and, again, the high heterogeneity regarding topics and research approaches. Besides, some of the reviewed studies did not compare tablet-based interventions with a non-technology control group. Still, Haßler et al (2016) concluded that there is little doubt that tablets can be viably used to support students learning.

However, data from large-scale observational studies have shown a different picture. Specifically, they indicate that the use of computers in education is even negatively correlated with students' achievement in various academic subjects (e.g, Hu, Gong, La, & Leung, 2018; OECD, 2015; Skryabin, Zhang, Liu, & Zhang, 2015), including reading performance (Hu et al., 2018; Judge, Puckett, & Bell 2006; Ponzo, 2011; Wenglinsky, 2005). They have also reported that a low or medium use of computers in classroom is indeed related to better learning outcomes than an intensive use, both in secondary students (Ponzo, 2011; OECD, 2015), and in higher education (Schmid et al., 2009). In this sense, the evidence highlights the influence of the quality of use of the educational technologies (Schmid et al., 2009; Tamim et al. 2011). Research have consistently found that digital technologies are generally underused and, most important, although teachers often have positive impressions about their usefulness as instructional tools (Eickelmann & Vennemann, 2017; Guha, 2003; Wang & Holthaus, 1999), they mostly use them for teacher-centered activities, such as presenting information through direct instruction, or for other basic activities, such as word processing or reinforcing learning skills through *drill and practice* activities (e.g., Becker, 2000; Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014). Indeed, the meta-analysis by Cheung and Slavin (2012) found that the types of computer-assisted reading interventions that have dominated the use of technology in the classroom during the past decades were those that yielded no additional benefits as compared to control groups. Bad news is that educational policies, practitioners, and even parents tend to assume that simply bringing digital technologies to educational institutions directly leads to a profitable use (e.g, Cuban, Kirkpatrick, & Peck, 2001; Pedró, 2012).

In short, evidence from research and international evaluations on the use of digital technologies as instructional tools shows that the learning outcomes are controversial and heterogeneous. Moreover, we should not ignore that results from large-scale data and meta-analyses on the effect of technology-based interventions are quite heterogeneus and inconclusive. However, whereas profitable applications and potential instructional uses of digital technology have received considerable attention (e.g, Cheung & Slavin, 2012, 2013; Haßler et al., 2016; Mulet, van de Leemput, & Amadieu, 2019; Spector, Merrill, Elen, & Bishop, 2014; Tamim et al, 2011), much less is known about cases in which the use of educational technologies yields poorer learning outcomes than traditional paper-based methods. Researchers, practitioners and policy makers should be aware of the challenges posed by the use of technology. To this end, we shall present in the Chapter 2 of the present dissertation a narrative review of empirical studies showing that digital tools are not always appropriate for instruction and learning.

Despite the considerable lack of attention to the possible hindering effects of technologies, our worries are not an exception, especially in the field of reading research. As mentioned above, concerns about the possible negative consequences caused by typical ways of accessing information through computers already arose when such devices started to infiltrate in workplaces, schools, and homes (Birkerts, 1994; Healy, 19989). Recently, some scholars have warned that the pervasive use of digital technologies could negatively affect cognition –e.g., sustained attention (Greenfield, 2015)– and, more specifically, the way in which we approach textual information, causing undesirable consequences in our ability to read in depth (Baron, 2015; Wolf, 2018). Whether digital medium affects reading processes and comprehension is the specific focus of the present dissertation. We next briefly review some of the most relevant theoretical models of reading comprehension, because, although they do not include the effect of media, their proposals illustrate the way in which reading on screen could affect reading comprehension. Subsequently, we shall review some empirical literature on the medium effect. Finally, the present Chapter 1 ends with some considerations regarding the main hypothesis on why and how on-screen reading affects reading performance.

1.2. Models of reading comprehension and the on-screen reading.

Traditional theoretical models of reading comprehension processes were developed to explain the understating of either written or oral discourse. Some of them propose explanations for comprehension of any type of information and even for the overarching cognitive activity. Although they do not take into account the possible influence of the reading medium (as we shall see next), a brief description of some of the most important models will help understand how reading on screen could affect reading performance. Firstly, we shall see that reading comprehension involves high-order cognitive processes (Kintsch, 1988, 1998) that require high level of mental effort. Secondly, considering reader-initiated processes as central for comprehension (Graesser 1994; van den Broek, Young, & Zeng, 1999) implies that reader-extrinsic factors that affect the reader could influence reading processes. Moreover, it is proposed that the reading context exerts a decisive influence on reading performance (Rouet, Britt, & Durik, 2017). As we shall see, the main hypothesis of the influence of the digital medium on reading is necessarily based on these three main assumptions, because it proposes that reading on screen hinders cognitive engagement (especially on-task attention and monitoring processes), causing a shallower comprehension of the text.

1.2.1. The Construction-Integration model

One of the most comprehensive and influential accounts for text comprehension is the Construction-Integration (CI) model. Developed by Kintsch (1988, 1998), this model is an ambitious theory aiming to explain not only discourse comprehension but also the overarching cognitive architecture (Wharton & Kintsch, 1991). Although considering the pivotal role played by prior knowledge (Kintsch, 1988), the CI model stresses bottom-up processes as responsible for text comprehension. The model was rooted in the initial approaches by Kintsch and van Dijk, who established three levels of representation of a text: the surface structure, the textbase, and the situation model (Kintsch and van Dijk, 1978; van Dijk and Kintsch, 1983). The simplest level is the *surface structure*, which represents the identification and short-term retention of words and the syntactic structures that connect words to each other (i.e., the verbatim forms). The next level of representational complexity is the *textbase*, which refers to the explicit semantic meaning and consists of a corpus of propositions. Propositions are basic processing units, each consisting of a complete idea or meaning unit that is represented in a simplified form that transcend the specific details of the surface structure. Finally, the highest level of comprehension is the *situation model*. It represents the reader's construction of a full representation of the text content, which is achieved by refining the textbase propositions and by blending them with prior knowledge. Thus, the situation model may vary across readers, because it depends on factors such as prior knowledge, interests, or task goals. "A situational model is the cognitive representation of the events, actions, persons, and in general the situation that a text is about" (van Dijk & Kintsch, 1983, pp. 11-12). It represents the deepest level of comprehension allowing long-term retention, so that "the concept of a situation model is required as a basis for learning" (p. 342), because it integrates the text into the reader's prior knowledge (Kintsch, 1998).

Upon the above assumptions, Kintsch developed his CI model, which consists of two ordered steps: *knowledge Construction* and *knowledge Integration*. According to him, the scripts or frames that guide text comprehension have to be effective enough to construct correct word meanings, propositions, and inferences (Kintsch, 1988). However, if they are too strong, they will not allow the reader to adapt to new discursive contexts. Thus, the construction process follows weak rules and it is prominently bottom-up and automated. Once the reader has identified and constructed and initial linguistic representation of the surface structure of the text, the Construction step entails the formation of a transient corpus of meaningful units ranging from wrong or incomplete propositions to proper propositions. Based on the linguistic representation of the text still available in memory, as well as on the reader's prior knowledge, the Construction step finishes once an elaborated propositional network is created. The elaborated propositional network is still an incoherent and meaningless text representation. The knowledge Integration step entails a process in which coherent propositions are selected, because they are enhanced by having more positive connections with other propositions among the discourse elements as they are more contextually relevant. In contrast, those propositions with fewer or negative connections with other information within the set of propositions are discarded. Thus, the two-step Construction-Integration process occurs iteratively across processing cycles, each cycle corresponding to the processing of an input segment (i.e., a sentence). Within each cycle new information is activated and some concepts from previous cycles are retained (Kintsch, 1998).

The concepts of textbase propositions and situational model reflect two different levels of comprehension: literal and inferential comprehension. On the one hand, literal comprehension consists of meaningful and coherent propositions that are explicitly mentioned in the text (i.e. the textbase). On the other hand, inferences go further by establishing connections between two different propositions not explicitly connected in the text –*bridging* inferences–, by connecting textbase propositions with the reader's prior knowledge –*associative* inferences–, or by producing new information derived from the information in the text –*transitive* and *logical* inferences– (see Kintsch, 1993, for an extensive classification of inferences). Thus, inferences are the product of higher-order comprehension, so they require an increased cognitive effort in which executive functions play a crucial role (Baddeley, 2003; van den Broek, Helder, & van Leijenhorst, 2013). Executive functions involve the coordination and combination of cognitive processes when performing complex cognitive tasks (Miyake et al., 2000) such as reading (Kendeou, McMaster, & Christ, 2016). Core components of executive functions such as working memory, planning, and inhibition processes have been found to be stronger predictors of inferential comprehension than of literal comprehension (Potocki, Sanchez, Ecalle, & Magnan, 2017). Therefore, deep comprehension of a text requires additional cognitive investment to construct the situation model, because it integrates all the inferences that the reader has made.

Finally, Kintsch highlights that both the textbase and the situation model are two dimensions of the reader's episodic memory of the text, so they should not be considered as separated representations. In fact, in extreme cases, the situational model could entirely consist of the textbase. As Kintsch (1998) argues, "in the text-that-tells-it-all, in which every detail as well as the overall structure is made perfectly explicit (as far as that is possible!), the textbase is also a good situational model and no further knowledge elaborations on the part of the comprehender are required" (p. 104). Nevertheless, he assumes that it is useful to distinguish these two components for analytic purposes in research and instruction. In this sense, the readers' construction of the textbase can be more or less coherent and complete, and the situation model (when it contains reader driven inferences, as it is often the case) can be more or less adequate and precise. As mentioned before, the more adequate and precise the construction of the situation model the higher the level of comprehension of the text content.

1.2.2. The Constructionist theory

Although the CI model proposes a comprehensive account for comprehension, it is prominently based on a bottom-up view of the reading process. It explains comprehension rather as a passive, memory-based activity than an active, reader-driven activity. Filling this gap, the Constructionist theory (Graesser, Singer, & Trabasso, 1994; Singer, Graesser, & Trabasso, 1994) presents reading comprehension as an active, reader-driven activity in which readers' goals, intentions, and strategies play a central role. The theory was founded on Bartlett's principle of *effort after meaning*, that refers to any cognitive activity that attempts "to connect something that is given with something other than itself" (1932/1995, p. 227). Also known as *search after meaning* in the specific field of comprehension (Stein & Trabasso, 1985), this principle entails three central assumptions that were explicitly adopted by the Constructionist theory: 1) the reader goal assumption, 2) the coherent assumption, and 3) the explanation assumption. Briefly, they altogether consider that "readers attempt to construct a meaningful referential situation model that addresses the readers' goals, that is coherent, and that explain why actions, events, and states are mentioned in the text" (Graesser et al., 1994, p. 372).

In order to achieve the above characterization of comprehension, readers generate inferences (this is the *search-after-meaning* principle). Thus, the main goal of the Constructionist theory is to predict how readers generate on-line (i.e., while reading) three specific types of inferences: the superordinate goals that motivate the action of the characters; the causal antecedents that explain why actions, events, or states are mention in the text; and global thematic inferences that construct the gist of the text. The readers often generate these inferences when needed, unless they think that the text is not well written so it has no global coherence, they have not enough background knowledge, or their goals do not demand constructing a coherent situational model (Graesser et al., 1994).

The Constructionist theory also describes six production rules that implement the assumptions of the model, so they generate the inferences. Each rule can be fired within each comprehension cycle depending on the explicit statement mentioned in the text. Thus, six different conditions that triggers the rules are described, which vary in the statements in the text (e.g., "Explicit statement in the discourse focus is an intentional action or goal of a character"), or in the activation of different elements in the reader's working memory (e.g., "An implicit statement or structure in working memory meets some activation threshold"; Graesser et al. 1994, Table 2). Each condition activates different cognitive processes that generate the inferences, involving: a) searching for information sources in long-term memory and working memory; b) searching for information within information sources; c) increasing the activation of content in working memory; and d) verifying whether potential inferences are compatible with the active content in the working memory (Graesser et al., 1994).

The main contribution of the Constructionist theory is that it accounts for the generation of inferences by means of processes that are initiated by the reader. Thus, it allows to identify the active processes that readers apply when learning from texts in educational contexts, which in turn have important implications regarding educational interventions aiming to improve, for example, reading strategies (Graesser, 2007). However, this could also be considered as their main limitation, given that the model is correct "when the reader is attempting to comprehend the text for enjoyment or mastery at a more leisurely pace, when the text has global coherence, and when the reader has some background knowledge" (Graesser, Mills, & Zwaan, 1997, p. 183).

1.2.3. The Landscape model

Together with the CI model (Kintsch, 1988, 1998), the Landscape model is the most comprehensive proposal to date, developed by van den Broek et al. (Linderholm, Virtue, Tzeng, & van den Broek, 2004, van den Broek & Helder, 2017; van den Broek et al., 1999). However, unlike the CI model, the Landscape model is based on the idea that neither bottom-up nor pure top-down processes should be considered in their pure form, because bottom-up processes often require top-down information to some extent –e.g., those concepts from prior knowledge that will be activated by the text–, and vice versa –e.g., when previous schemas interact with the textual meaning (van den Broek, Rapp, & Kendeou, 2005)–. According to the authors, although previous models of comprehension, such as those described above, included both bottom-up and top-down processes as responsible for comprehension to some degree, they do not address how these two types of processes relate and combine. Thus, filling this gap is one of the main goals of the

Landscape model.

The name of the model comes from the idea that, due to the processing constraints of cognition, the reader only can pay attention to a set of the concepts conveyed by the text and the relations between them at the same time. Thus, the activation of these elements in reader's mind as he or she go forward through the text is fluctuant. The result is a "landscape of fluctuant activations" that are not deemed as all-or-none activations but as gradients (van den Broek et al., 1999, p. 73). Based on this assumption, the model was developed by considering four sources of activation: 1) the current text segment –or reading cycle (Kintsch, 1988)–; 2) the information activated in the immediately preceding cycle, which is still available to some extent; 3) the information reinstated from a previous reading cycle; and 4) the reader's background knowledge. The central idea is that the fluctuating activation is not merely text-driven, but it is the result of the interaction between the text, the reader's cognitive capacities and prior knowledge, the mental representation of the text being created, and the readers' decisions through the comprehension process, that can be automatic or strategic (van den Broek et al., 1999; van den Broek & Helder, 2017). These four sources are deemed to interact within the reading process in order to activate concepts and to create associations between concepts, resulting in the construction of the mental representation of the text. However, the authors considered that the first source of activation is self-evident, because it is the text segment being read, and that the second source has been already addressed by previous models – e.g., the CI Model (Kitsch, 1988)–. Thus, the Landscape model is rather focused on how readers reinstate information from previous reading cycles, and on how they activate information from background knowledge (Linderholm et al., 2004).

Two mechanisms drive the reinstatements of information and the activation of prior knowledge: *cohort activation* and *coherence-based retrieval*. The mechanism of cohort activation is passive and memory-based, because it is triggered by the text elements. This mechanism activates concepts that, in turn, activate other associated concepts from the reader's prior knowledge or from his/her mental representation of the text (i.e., from the content of previous text segments). The cohort activation mechanism is similar to the knowledge Construction process proposed by the CI model (Kintsch 1988, 1998). On the other hand, the mechanism of coherence-based retrieval is reader-initiated and represents a strategic behaviour by which relevant information is retrieved also both from prior parts of the text that are already part of the reader's mental representation and from prior knowledge. This reader-initiated mechanism is similar to the search-for-meaning process proposed by the Constructionist theory of comprehension (Graesser et al., 1994), and it requires the reader to invest additional cognitive efforts, although they can be routinized to some degree (van den Broek & Helder, 2017)

Reader's coherence-based decisions rest on the concept of reader's *standards of coherence*, which refers to the desired level of coherence (e.g., causal, referential coherence) that the reader aims to reach when constructing the mental representation of the text. In other words, the standards of coherence represent what the reader considers an adequate comprehension of the text at hand. Passive processes are always triggered by each new text segment; however, reader-initiated processes will not occur unless the reader needs to increase the level of comprehension to achieve his or her standard of coherence, which in turn depends on factors such as the text characteristics, task goals, motivation, cognitive capacities, or background knowledge (van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011). Therefore, if passive processes provide a mental representation that reach reader's standard of coherence, he or she will not invest efforts on additional strategic processing. Reading comprehension is hence the result of the reciprocal interaction between passive processes, reader-initiated processes, and the evolving mental representation of the text content. This interaction fluctuates cyclically over the text segments, with both types of processes working in parallel (van den Broek & Helder, 2017) and depending on reader's limited attention and working memory (van den Broek, 2010).

The Landscape model is considered to be the most comprehensive model or reading comprehension mainly due to its emphasis on integrating memory-based and on constructionist views of reading (e.g., van den Broek et al., 2005). Furthermore, given that most comprehension models have been developed based on reading of narrative texts (McNamara & Magliano, 2009), an additional relevant contribution by van den Broek and colleagues is that the model has been applied also to expository texts (e.g., van den Broek & Kendeou, 2008). Learning from this type of texts require to comprehend associations between facts, concepts, or processes, in order to integrate them into reader's prior knowledge. For these purposes, simultaneous activation of different pieces of information often far apart within the text is fundamental. Within this scenario, given the limitations of reader's working memory, the allocation of attention to relevant informational elements becomes especially crucial (van den Broek, 2010).

1.2.4. The RESOLV model and the reading context

Taking into account reader-initiated processes as responsible for the reading comprehension allows to consider reader-extrinsic factors others than the text that could affect the reader, hence the reading process. Thus, there is a burgeoning research corpus over the past years focusing on how the contextual factors influence reading performance

–e.g., the task instructions (McCrudden, Magliano, & Schraw, 2010), the influence of peers or tutors when learning from multiple texts (Rouet & Britt, 2011), or even overarching circumstances such as social group membership (Snow, 2002)–. As expressed by Rouet and Britt (2011), "given a particular reader and a particular text, many different reading behaviours may be observed as a function of when, where, and why the reading episode is taking place" (p. 21). Based on this assumption, Rouet, Britt, and Durik (2017) have recently proposed the RESOLV model, which considers the reading context as a crucial influence for reading performance.

The RESOLV model lies on the idea that reading is an activity performed within a physical and social context that determines conditions and resources for reading. Accordingly, the reader's comprehension of the text is the product of the interaction between him or her context model, task model, and reading processes. The model proposes that "reader's processing of contextual features always play a part in their engagement with text", and that "the construction of a context model is triggered by readers' attention to, perception of, and interpretation of the physical and social situation" (Rouet et al., 2017, p. 206). Thus, it states that the readers construct their context representation based on features such as people, place, time, objects, and tools. Importantly, the model assumes that readers construct and store in memory a schema of each typical reading situation they have previously experienced, and that they do not perform a systematic evaluation of all the environmental factors to activate the context schema. Instead, the readers process only a subset of the contextual clues that is informative enough to trigger a particular schema. To do this, they rely on their prior experience of similar situations. Importantly, each context schema involves a set of reading demands, purposes, and goals, and a set of reading actions and strategies to adapt to these requirements (e.g., reading for comprehension, skimming, skipping, establishing comprehension standards). Thus, besides the influence of task demands, readers construct a representation of the context consisting of a range of reading actions on which they based their reading decisions.

Rouet et al. (2017) mention different subject classrooms as examples of different reading contexts. The readers activate different reading purposes and actions depending on the subject (e.g., language vs. science). Another example of a typical context is a situation in which a reader look for information on the Internet as part of an informal discussion (Rouet et al., 2017). In this sense, if one wonders about the diversity of contexts in which nowadays reading takes place, the digital environment quickly comes to mind.

Digital devices and the Internet represents a new substrate for reading which potentially could even replace printed texts as the dominant medium (Mangen & van der Weel, 2016), and it is widely assumed to be qualitatively different to our traditional way of reading. In Mangen and van der Weel's (2016) words, "digitisation has dislodged reading from its natural place in the constellation of modalities and media. The static, linear modality of written text (including the book) is now supplemented by an increasing complexity of multimodal, dynamic, and interactive representations" (p. 116). Thus, there is a vast literature corpus focused on the necessary abilities and skills to cope with a complex set of activities such as searching, navigating, accessing multiple documents, evaluating sources, and manipulating information in a non-linear form (e.g., Alexander & The Disciplined Reading and Learning Research Laboratory, 2012; Coiro & Dobler, 2007; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018; Segers, 2017; Wylie et al., 2018).

On the other hand, scholars are also interested in whether readers comprehend as good when reading on screens as when reading printed texts. In this regard, three main questions arose: 1) why should we expect that reading on screen affects reading performance?; 2) does reading on screen effectively affect reading comprehension?; and finally, if the answer to the second question is (or could be) yes, 3) what is the plausible explanation for this effect? Based on these questions, we shall firstly review some findings regarding readers' preferences, impressions, and habits when using digital technologies with learning purposes and, specifically, for reading. According to the RESOLV model (Rouet el al., 2017), the experiences that people usually have when reading on screen would construct a context schema that activate a specific reading setting. Thus, if reading on screen involves a particular reading context, it will in turn affect the reading process and, hence, the outcomes. Secondly, we shall briefly review some literature comparing reading comprehension outcomes between reading these reading media. Finally, we shall present the Shallowing hypothesis as an explanation for the hindering effect of the digital medium on reading comprehension.

1.3. Readers' medium preferences

The history of printed texts in modern education is a 100-year history of resistance. In 1913, Thomas Edison predicted that the motion picture would dominate instructional methods. He said in an interview that "books will soon be obsolete in the public schools" (cit. in Morris, 2019, p. 146). Almost 100 years later, Jeff Jarvis wrote that "thanks to the searchable, connected Internet, books could be so much more […] While we worship the book with its present limitations, we cannot reinvent it. The book is dead. Long live books" (Jarvis, 2006). Nevertheless, more than 10 years after Jarvis's prediction, the vast majority of the studies exploring today's students' preferences and impressions regarding the use of electronic texts consistently find that they still prefer printed texts, especially when it comes to reading in depth and to learning from texts.

Still, digital texts are not rejected at all. Students usually report that they prefer electronic documents when they have to look for and access different sources or pieces of information. For example, Liew, Foo, and Chennupati (2000) found that 73.5% of a sample of 67 graduate students preferred reading electronic scientific journals rather than their printed versions. Main reasons were that they provide hyperlinks to additional information, easy and quick access, and the currency of documents. However, in spite of this preference, they mostly consider reading on paper what they found on-line as the best option, and 83.1% percent agreed that it is difficult to read from screen. Similarly, Sathe, Grady, and Giuse (2002) surveyed medical and nursing students, researchers and professionals about their use of medical journals. Only 17% of the participants declared they preferred to use print journal, whereas 50% indicated their preference for using them in the electronic format. Similarly, this group mentioned ease of access and searching as the main reason for using the electronic format. Paradoxically, the second most-mentioned reason was the ease of printing (Sathe et al., 2002). Liu (2006) consistently found that although 84.2% of a sample of 133 undergraduates declared to use electronic journals rather than printed journals, 81% indicated that they print out the electronic documents 'all the time' or 'most of the time' that they have to read them.

With respect to the use of electronic textbooks (e-textbooks), Woody, Daniel, and Baker (2010) asked 91 undergraduates to answer some questions on satisfaction and habits regarding the use of this new type of text-books. They found that those that had used an e-textbook in a previous course were less satisfied with it than those who used a printed textbook. Additionaly, participants declared to be more likely to read captions and charts in printed textbooks than in e-textbooks. Finally, they indicated a moderate preference to learn using a printed textbook regardless of they had used an e-textbook before (Woody et al, 2010) –for similar findings on undergraduates' preference for printed textbooks see Buzzetto-More, Guy, & Elobaid (2007), Chulkov & VanAlstine (2013), Daniel & Woody (2013), and Shepperd, Grace, & Koch (2008)–. But some positive attitudes toward the use of e-textbooks have also been described. For example, Dobler (2015) found that although only 22% of a sample of 56 preservice teachers preferred to use an e-textbook for a methods course, the percentage increased to 50% after using it during the course, although a 42% of participants still were reluctant. Interestingly, participants highlighted the role of the teacher in helping them to properly use the e-textbook and its affordances (Dobler, 2015). Moreover, a recent study found similar satisfaction for both textbook formats in 42 undergraduate students that used a printed textbook and an e-textbook in each semester, respectively, within an academic year (Sommers, Shin, Greenebaum, Merker, and Sanders, 2019). The participants also showed a more positive attitude with respect to using the e-textbooks after experiencing this modality. Moreover, they declared to engage more often with the end-of-chapter multiple-choice questions when using the e-textbook, probably because this version of the questions gave instant feedback on the response, authors argued (Sommers et al., 2019).

Nevertheless, recent studies have also found an overarching trend of preference toward reading in print, especially when it is required a high level of engagement with the text. Therefore, it seems that readers' impressions on this regard have not changed over the last 15 years. Notably, a large-scale survey conducted by Mizrachi, Salaz, Kurbanoglu, Boustany, and the ARFIS Research Group (2018) covering more than 10,000 college students from 21 countries found that 78.4% of the respondents preferred print format when asked about reading their academic course materials. Similarly, Kurata, Ishita, Miyata, and Minami (2017) explored reading habits in a sample of 1,755 adults, ranging from 18 to 69 years old. Results indicated that, despite the participants reported that approximately 70% of their daily reading time was spent reading on digital devices, the majority preferred reading in print, both for working purposes (80.2%) and for leisure reading (76.35%). Furthermore, it seems that the text length matter. A study conducted by Baron, Calixte, and Havewala (2017) found that whereas 42.6% of a 429-undergraduate sample declared preferring the hard copy when reading short texts, the percentage increased to 86.4% when they were asked about reading long academic texts –for similar findings see Farinosi, Lim, & Roll (2016), Mizrachi et al. (2018) and Stoop, Kreutzer, & Kircz (2013).

1.4. Readers' habits when using digital technologies

As mentioned above, in spite of readers' preferences towards reading in print, the digital medium is not banned. Readers prefer electronic texts and reading on-line when they aim to access different sources or look for information (e.g., Baron et al., 2017; Liu, 2006; Rose, 2011; Stoop et al., 2013). Additional positive views are related to the fact that electronic texts are searchable and cost-saving (e.g., Ji, Michaels, & Waterman, 2014; Mizrachi et al., 2018; Muir & Hawes, 2013). Nonetheless, observational studies have consistently found that a large share of today's young adults still prefers to read printed texts.

The question is, why people still prefer reading on paper? Exploratory studies have consistently found that people mostly report that reading on screen prevent them to fully concentrate and to engage with the reading task. In other words, they usually indicate that paper facilitates text immersion and in-depth reading (e.g., Baron, 2013; Baron et al. 2017; Dobler, 2015; Farinosi et al., 2016; Liu, 2005; Mizrachi et. al, 2018; Rose, 2011; Potnis, Deosthali, & Pino; 2017). Moreover, people often indicate that they behave differently when reading on screen with respect to actions such as marking, annotating, or highlighting, that seem to decrease in this medium. For instance, 53.9% of the respondents in Liu (2005) reported that they frequently annotate while reading on paper, whereas only 10.7% declared to do it when reading electronic documents. In general, readers indicate that the print format fosters actions that implies effortful reading and favour in-depth comprehension (Baron, 2013; Foasberg, 2014; Ji et al., 2014; Mizrachi et al., 2018). In this direction, an experimental study comparing reading performance found that readers in the on-print reading group annotated more often than those in the on-screen group (Norman & Furnes, 2016). It has also been reported that, except for highlighting, most students do not used mark-up tools when utilizing a e-textbook for a one-semester university course, although they were taught how to use them (Van Horne, Russell, & Schuh, 2016). Moreover, readers indicate that activities such as scanning and one-time reading increase when reading on screen (Liu, 2005), as well as eye-tracking studies have found that skim reading and scanning are the usual ways of reading when gathering information on the Internet (Pernice, Whitenton, & Nielsen, 2014).

Digital devices are commonly considered in literature highly distracting stimuli, a view based not only on respondents' perceptions about their own attentional status but also on how people usually interact with computers, tablets, or smartphones. For example, exploring the frequency of students' multitasking when using digital technologies, and its effects on learning is a recurrent issue in educational research, and there is a strong consensus about that digital devices foster multitasking. For example, Levine, Waite, and Bowman (2007) and Baron et al. (2017) found that more than 60% of the participants in their surveys admitted they multitask while working and reading on computers. Furthermore, Levine et al. (2007) found that the amount of use of instant messaging predicted the level of self-reported distraction during academic reading[1]. In this line, Daniel and Woody (2013) conducted an experimental study in which some participants were asked to read a textbook chapter in print or in electronic formats at home. After the reading task, participants rated how often they engaged with competing activities while reading. Findings revealed that those who read in the electronic formats engaged more often with activities such as messaging, e-mailing, accessing social media, and interacting with other people at home (e.g., roommates, relatives).

Media multitasking –i.e., "a person's consumption of more than one item or stream of content at the same time" (Ophir, Nass, & Wagner, 2009, p. 15,583)– is a pervasive activity, especially in younger generations (Baron, 2013; Daniel & Woody, 2013; Greenfield, 2009; Rideout, Foehr, & Roberts, 2010; Rideout, 2015; Rosen, Carrier, & Cheever, 2013). Rideout et al. (2010) reported that 29% of the time that 8-to-18-years-old students used media, they used two or more media at the same time. Moreover, 27% of the participants declared to use another media 'most of the time' when reading. Rosen et al. (2013) observed that students constantly switched attention to social media and texted to messaging on computers while studying at home. Furthermore, it has been found that most of the students who declared that they often multitask while doing their

[1] Participants rated statements such as "I feel impatient while reading my textbooks".

homework believed that this behaviour did not affect negatively their work (Rideout, 2015). Nonetheless, it has been consistently reported that multi-tasking hinders cognitive processing during task execution (Bergen, Grimes, & Potter, 2005; Foerde, Knowlton, & Poldrack, 2006; Koch, Lawo, Fels, & Vorländer, 2011). Daily multitasking activity has been also found to correlate negatively with some cognitive abilities such as executive control (Sanbonmatsu, Strayer, Medeiros-Ward, Watson, 2013; van der Schuur, Baumgartner, Sumter, & Valkenburg, 2015) and the ability to filter irrelevant stimuli from the environment or from one's memory (Ophir, Nass, Wagner, 2009; but see Wiradhany & Nieuwenstein, 2017), as well as with academic achievement (Carrier, Rosen, Cheever & Lim, 2015; Junco & Cotton, 2012; Fox, Rosen, & Crawford, 2009; Rosen et al., 2013; van der Schuur et. al, 2015).

The hindering effect of the so-called *technological distraction* on learning has been experimentally studied by examining the effect of introducing digital devices (mostly laptop computers) in classroom during lectures or while studying at home. For example, Sana, Weston and Cepeda (2013) conducted an interesting study on the effects of on-line multitasking during a university lecture. In Experiment 1, they found that those who were asked to complete some extraneous tasks on their laptops during the lecture (e.g., checking what was on a TV channel at 10 pm) scored lower in a post-lecture test about the lecture content (see also Wood et al., 2012 for similar results). Furthermore, Hembrooke and Gay (2003) found that even when students were encouraged to use their laptops as a learning supplement during lectures, they scored significantly lower in a recall test than those who kept their laptops closed, probably because they also engaged in off-task on-line activities such as web-browsing or accessing social media. In this regard, Kraushaar and Novak (2010) monitored students' laptop activity by means of a spyware, and they found that 42% of time that students engaged on multitasking during lectures was on non-course related activities. Importantly, even when laptops are used by nearby peers in classroom during lectures, they are distracting stimuli that negatively affect those that did not use them (Sana et al., 2013, Experiment 2). Similarly, surveyed students in Fried (2008) indicated that fellow's use of laptops represented for them a major source of distraction during lectures.

The studies presented above highlight that multitasking is an activity that people massively practice when using digital technologies on a daily basis. Hwang, Kim and Jeong (2014) explored the motives for media multitasking in adults ranging from 19 to 59 years old and they found that this activity is often performed as a mere habit. Media multitasking is thus something that people usually simply do when they face digital technology, even while doing activities that demand cognitive effort and focusing, such as reading or doing homework (Baron, 2013; Daniel & Woody, 2013; Rideout et al., 2010; Rideout, 2015; Rosen et al. 2013). Moreover, digital devices seem to be external stimuli that prevent learners' on-task attention even when they are used by nearby fellows. Thus, students' impressions that reading on screen is detrimental for cognitive engagement and attention could be based on their media multitasking habits, which often lead them to non-task-related activities. Therefore, they perceive digital technology as inherently distracting devices that prompt to involuntary shifts of attention from task-related activities (Fried, 2008). Findings from Rosen et al. (2013) illustrates the problem. They found that college students who were observed while studying at home were, on average, less than 6 minutes focused on task before switching to computers to access social media and to do texting. This "*habitual distraction*" has been characterized as a consequence of people's "deeply sedimented relational strategies" that shape how we interact with the environment when digital technology is on hand (Aagaard, 2015, p. 95).

In short, digital technologies could be external factors that prompt a cognitive status characterized by constant attentional shifting, so that they prevent people to fully engage with the task at hand. Primary observational and experimental studies have shown that media multi-tasking is massively practiced by people when using computers, and that it often implies engaging with off-task activities. Accordingly, young adults mostly indicate that reading on screen prevents them from fully concentrating and engaging with the text. They thus prefer the printed medium, especially when reading long texts and with learning purposes. Based on these circumstances, the next question is: does the medium matter when it comes to comprehending and learning from texts? We shall next briefly review what literature on this issue has revealed during the past decades.

1.5. Empirical evidence of the medium effect on text comprehension

Some of the first experimental studies examining differences between reading on screen and reading on paper were reviewed in a seminal work conducted by Dillon (1992). The author concluded no apparent differences between both media with respect to reading comprehension outcomes. However, he noted that reading on screens seemed to be slower and sometimes less accurate than on-print reading. Dillon also highlighted that the studies were highly heterogeneous, so that it made it difficult to reach any solid conclusion. Yet, he argued that differences, when appeared, seemed to be mainly due to ergonomic factors, especially to the image quality, but he emphasized that single variable explanations cannot not successfully address the issue. Dillon's work was quite remarkable, but it suffered from several limitations. It was not systematic, since he did not describe the search procedure or explicitly reported how the reviewed studies were selected.

Moreover, his conclusions were mostly based on single studies instead of on a compilation of studies (for an extensive critic see Singer & Alexander, 2017). Still, this review entailed a starting point for a growing line of research.

As mentioned above, the use of computers in educational settings have increased since Dillon published his review. Accordingly, the number of studies focus on their influence on reading have also grown, mostly focused on reading comprehension outcomes. Over the last years of the 90's and the first years of the 21st century, as standardized reading tests and official evaluations of students' reading achievement began to be designed and applied in computerized form, research have been especially focused on examining their validity, as compared to their traditional printed versions. Two meta-analyses conducted by Wang, Jiao, Young, Brooks, and Olson (2007) and Kingston (2008), respectively, synthesized these first media studies. The meta-analysis by Wang et al. (2007) compiled findings from studies on reading testing in K-12 population that were published since 1988 to 2005 (although most of them were conducted in 2004). In total, the dataset included 42 effect sizes from 11 studies, which were conducted both with tests whose computerized version was adaptive to test-taker's level, and tests with a linear (i.e., fixed, non-adaptive) computerized version. In general, Wang et al. found a small but significant effect of the presentation mode favouring the printed application (ES = 0.08). However, the authors noted that given the size of the mean effect, according to Cohen's (1988) classification, the advantage for the paper-based tests was negligible. Interestingly, the analysis of moderators indicated that the advantage of paper-based testing was larger in those tests whose computerized version was linear, indicating that the more comparable the versions the larger the advantage for paper. Unfortunately, the value of the effect size for each of these different sets of studies was not reported.

For its part, the meta-analysis by Kingston (2008) compiled studies also conducted in K-12 populations. This work was focused not only on reading but also on other academic subjects including mathematics, language arts, social studies, and science. Although the publishing dates ranged from 1997 to 2007, all the studies but one were published in 2002 or later. Contrary to Wang et al. (2007), the results from Kingston's meta-analysis yielded, on the one hand, no difference between media among the studies examining reading achievement (ES = –0.01, including 30 effect sizes from eight studies). On the other hand, there was a small but significant effect size (0.11) favouring computerized mode among the studies that assessed achievement in language arts (nine effect sizes from four studies). In sum, the main conclusion that can be drawn from these two meta-analysis is that these first studies comparing reading comprehension across reading media yielded inconclusive results, as well as that the differences, when appeared, were relatively small. In addition, a clear limitation of these two meta-analyses is that they included a small sample of studies. This circumstance, together with the heterogeneity of their results, makes it hard to draw any strong conclusion.

Experimental studies on the effect of reading media on comprehension have been conducted especially over the past 15 years. As we shall see below, although most of them were conducted using computers, research is paying an increasing attention to the use of hand-held devices for reading. Besides the use of different digital devices, research is highly heterogeneous for other relevant methodological characteristics such us participants' age, texts genre, texts length, the type of questions (e.g., literal questions, inferential questions, recall, recognition, multiple-choice), or reading time constraints. The studies that used digital texts that include features and affordances that are digital-exclusive (e.g., audio-visual content, hypertext, Internet navigation) are out of the scope of this dissertation, because we aim to examine the effect of reading on screen per se. Thus, we shall not take into account this type of investigations. We next present only a brief overview of the state of the art, because a systematic and exhaustive analysis of the literature in this field can be found in the meta-analysis included in Chapter 3.

Not surprisingly, the medium effect does not appear consistently across the studies on this issue. On the one hand, not a few works have reported essentially similar results in comprehension outcomes across reading media. For example, this absence of differences was found either in elementary-school students who read excerpts of a storybook (Grimshaw, Dungworth, McKnight, & Morris, 2007) or a mix of narrative and expository passages (Higgins, Russell, & Hoffmann, 2005), in high-school students who read an expository text (Porion, Aparicio, Megalakaki, Robert, & Baccino, 2016), and in undergraduates who read expository texts (Mayes, Sims, & Koonce, 2001; Taylor, 2011) or narrative texts (Seehafer, 2014). Nevertheless, a remarkable body of studies found that reading on computer yielded poorer reading outcomes, for instance, in elementary students who read a mix of narrative and expository texts (Lenhard, Schroeders, & Lenhard, A. (2017) and narrative texts (Jeong, 2012), in high-school students who read an expository text (2013), as well as in undergraduates who read expository texts (Singer & Alexander, 2017b), or a combination of a narrative and an expository text (Mangen, Walgermo, & Brønnick, 2013).

Additionally, other studies have also found poorer comprehension outcomes of reading on screen but under specific circumstances. For example, in Ackerman and Goldsmith (2011) the on-screen reading inferiority arose when the participants self-paced their reading time –vs. pressured reading time (participants were given approximately

70% of the time invested by the self-paced time group)–. In contrast, in a follow-up study this effect appeared when the participants read under time pressure (Ackerman & Lauterman, 2012). Authors argued that this inconsistency between the time-frame and the medium effect across both studies was due to differences between each sample of participants. Furthermore, increased difference in comprehension favouring reading in print have also been found when participants were asked to use annotation tools while reading –vs. reading without annotating (Ben-Yehudah, & Eshet-Alkalai, 2014)–, as well as when readers had to scroll down the texts when reading on the computer –vs. a page-by-page presentation (Higgins, Russell, & Hoffmann, 2005)–, although in this latter case the effect was not significant.

Finally, another possible source of variation could be the digital device use for reading on screen. As mentioned above, most studies to date have been conducted with computers. However, the interest in hand-held devices (mostly tablets) has increased during recent years as they become more popular[2] (Mulet et al., 2019). As can be seen in Table A1 and Table A2 in Appendix A, studies with these devices represent approximately half of the research works conducted from 2012 to 2017. Once again, if we look at this set of studies and any other recently published, we find the same heterogeneous results. Whereas some of them found no differences between media (e.g., Kretzschmar et al., 2013; Hermena et al., 2017; Sackstein, Spark, & Jenkins, 2015), others reported that reading on paper yielded a higher frequency of high scores in comprehension (although mean scores were similar; Niccoli, 2015), as well as better results in comprehension, both in eighth graders (Simian et al., 2016), and in

[2] Note that Apple's iPad (the first tablet) was first released in late 2010.

undergraduates (Haddock, Foad, Saul, Brown, & Thompson, 2019; Nishizaki, 2015, Experiments 1 & 2). Moreover, two studies reported opposite comprehension results favouring on-tablet reading in elementary-school students (Lian & Huang, 2014, Task 3; Nishizaki, 2015, Experiment 1). In addition, Chen, Cheng, Chang, Zheng, and Huang (2014) compared reading on paper with reading on tablet and on computer. Their results indicated that although the on-print readers performed better on a reading comprehension test that the on-screen groups, this effect was significant only compared to those who read on computers. On-tablet readers performed slightly better than on-computer readers and worse than on-print readers, but these differences were not significant.

Two main conclusions can be drawn from the findings mentioned above. Firstly, the research body is highly heterogeneous, not only with respect to methodological features but specially with regards to the existing evidence of the medium effect. Secondly, although this effect does not appear consistently, it seems that reading on screen could be detrimental for reading comprehension, at least under some circumstances. Furthermore, when the medium effect arises, it consistently favours reading printed texts. Similar conclusions were drawn by Singer and Alexander (2017) after reviewing within-participants studies on the medium effect (i.e., all participants read on both media) from 1992 to 2017. Thus, given the inconclusive results and the methodological heterogeneity of the literature in this field, we decided to carry on a meta-analysis aiming to figure out whether the reading medium influences reading comprehension, and to know whether there are variables that qualify this effect. The results are reported in Chapter 3. Next, we shall present the Shallowing hypothesis as the main attempt to explain why reading on screen could hinder reading performance and text comprehension.

1.6. The shallow on-screen reading

The Shallowing hypothesis is based on the idea that readers fail to read in depth when they read texts on screen, resulting in a shallower processing of the information. This hypothesis is consistent with readers' impressions about on-screen reading, as well as with findings about the habitual shifting attention when performing learning activities with digital devices. Furthermore, it also lies on the assumption that digital reading is characterized by quick interactions with short pieces of information that shapes a reading habit that is incompatible with the required mental efforts to deeply engage with texts. Commentators such as Carr (2010), Baron (2015), or Wolf (2018) have raised well-funded concerns about this circumstance.

As mentioned above, people often indicate that digital reading habits usually consist in reading actions such as skimming, scanning and web browsing (Liu, 2005), and eye-tracking studies have confirmed such a superficial way of reading (Pernice et al., 2014). As described by van der Weel (2011):

> "The way these digital texts are consumed is very different too. Once networked, their full text can be searched as a body. This new form of access replaces the identification, location, and searching of relevant texts through the conventional bibliographical mechanisms that reigned in the world of print and imposed a hierarchical order on them. It brings novel ways of finding, promoting serendipity, but it also stimulates a sampling and zapping manner of reading. This way of consuming text is not unlike the way image and sound are consumed in today's world of multi-channel television and the seas of 'songs' that have replaced the 'albums' of yore." (p. 195)

Worries about this new type of *zapping* attitude to text as detrimental habit for sustained attention and for engaging with long texts have been really well expressed by Carr (2010), who have claimed that digital environments (i.e., the Internet) encourages and rewards a quick and superficial reading of the information. In this sense, it can be said that texts are not read but consumed. The idea that digital reading promotes a processing style opposite to what a deep understating demands has also been highlighted by Baron (2015) and Wolf (2018). Thus, the *zapping* attitude to text when reading on screen, together with the pervasive media-multitasking described above, would involve a broad range of reading habits that may create a mind-setting that would prevent readers to assign the mental efforts required to fully comprehend the textual information. This hypothesis, popularized by Carr (2010) but already proposed 15 years before by Bikerts (1995), has come to be known as the Shallowing hypothesis (Annisette & Lafreniere, 2017). In this sense, if we lay on the RESOLV model (Britt et al., 2017), we can argue that the digital device becomes a contextual cue that would activate a schema from readers' memory that, as a product of the usual type of interactions between readers and digital technology, is characterized by a lower level of on-task cognitive engagement and a shallower processing of the information.

In spite of the increasing interest in the medium effect on reading comprehension, there is a lack of research that have examined the underlying cognitive processes. In this regard, a series of studies conducted by Ackerman and colleagues have tried to demonstrate that the on-screen reading inferiority could be due to difficulties in monitoring one's level of performance when reading on this medium (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). The activation of metacognitive processes is considered a core cognitive activity when performing complex tasks, because they guide the allocation of mental effort and cognitive resources (Fiedler, Ackerman, & Scarampi, 2019). Thus, if reading on screen hinders the readers' metacognitive abilities, such as monitoring, it would support the Shallowing hypothesis. In this direction, Ackerman and colleagues concluded that people tend to overestimate their own level of comprehension and learning to a greater extent when reading on screen than when reading printed texts, and especially when reading under time pressure. Therefore, people would wrongly think that they have reached the desirable level of comprehension, which in turn would prevent them to invest the necessary mental effort to fully comprehend the text and to learn from it. Nevertheless, it is noteworthy that other studies did not find this effect (Chen & Catrambone, 2015; Norman & Furnes, 2016).

Three additional studies have also studied differences in cognitive engagement and efforts across reading media by measuring participants' eye movements (Bansi et al., 2016; Kretzschmar et al., 2013; Latini, Bråten, & Salmerón, 2019) and also EEG activity (Kretzschmar et al., 2013) while reading. Given that we shall discuss more in detail these studies in the introduction of Chapter 4, for the sake of avoiding redundancy, we shall only briefly mention their results here. On the one hand, Kretzschmar et al. (2013) found that, overall, cognitive efforts, as measured by readers' eye movements and EEG activity, did not differ between reading media. On the other hand, eye-movement data in Bansi et al. (2016) showed longer fixation times and shorter saccades when reading on paper, although they did not find differences in comprehension between reading media. Moreover, Latini et al. (2019) showed that readers in the on-print condition did more gaze transitions between pictorial and textual information when reading an illustrated text. Thus, readers in the on-print condition showed a better comprehension of the text. The authors argued that integrative processing was higher when reading in print. In short, whereas evidence from Kretzschmar et al. (2013) did not support the Shallowing hypothesis, findings from Bansi et al. (2016) and Latini et al. (2019) seems to support it.

To date, evidence that allows to test the Shallowing hypothesis is scarce and inconclusive, so much more research is needed on this issue. To that end, we present on Chapter 4 and Chapter 5 two experimental studies aiming to examine the circumstances that could qualify the possible effect of reading medium on reading comprehension, and whether the Shallowing hypothesis explains this effect. Firstly, in Chapter 4, we present a study conducted in a sample of 116 undergraduate students that were asked to read several expository texts either in print and on a tablet. We tried to replicate findings by Ackerman and colleagues (e.g., Ackerman & Lauterman, 2012) revealing a deficit in the metacognitive monitoring of comprehension when reading on screen under time pressure. We also measured participants' eye movements while reading in order to compare cognitive engagement in both media. Secondly, in the study presented in Chapter 5, 140 undergraduates were asked to read a long expository text on the actual printed magazine or on a computer screen. We also explored participants monitoring of their level of comprehension. Additionally, we tested the Shallowing hypothesis by measuring participants' on-task attention with mindwandering probes. As discussed more in detail in Chapter 5, mindwandering represents self-generated thoughts that are not related to the task currently performed (Smallwood, 2013). Thus, the level of mindwandering is considered an indicator of on-task attention, so of cognitive engagement with the task. To date, this study represents the first experimental attempt to compare on-task attention between reading on paper and reading on screen. Additionally, we shall examine the possible qualifying effect of the reading time-frame (i.e., self-paced vs. constrained reading time) on the medium effect on reading performance.

To summarize what the reader will find next, we shall firstly present in Chapter 2 a narrative review of previous empirical studies showing that digital technologies are not always suitable for learning. Secondly, Chapter 3 consists of a meta-analysis of the existing primary studies empirically comparing reading comprehension outcomes between reading on paper and reading on screen. Thirdly, as mentioned in the previous paragraph, Chapter 4 and Chapter 5 present two experimental studies aiming to further explore whether reading on paper promotes a better comprehension than reading on screen; to replicate findings showing difficulties in metacognitive monitoring when reading on screen (e.g., Ackerman & Lauterman, 2012); and to test the Shallowing hypothesis as an explanation for the on-screen inferiority by measuring participants cognitive engagement through their eye movements (Chapter 4) and through their level of on-task attention (Chapter 5). Finally, in Chapter 6 we shall discuss about the general conclusions and educational implications of this dissertation, and we shall suggest future research on the reading medium effect.

References

Aagaard, J. (2015). Drawn to distraction: A qualitative study of off-task use of educational technology. *Computers & Education*, *87*, 90-97. https://doi.org/10.1016/j.compedu.2015.03.010

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied, 17*, 18-32. https://doi.org/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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Alexander P. A., & The Disciplined Reading and Learning Research Laboratory (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist, 47*, 259-280. https://doi.org/10.1080/00461520.2012.722511

Anderson, R. E., & Ronnkvist, A. (1999). The presence of computers in American schools. Irvine, CA: Center for Research on Information Technology and Organizations, University of California, Irvine.

Annisette, L.E., & Lafreniere, K.D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154-158. https://doi.org/10.1016/j.paid.2016.02.043

Baddeley, A. (2003). Working memory and language: An overview. Journal of

Communication Disorders, 36(3), 189-208. https://doi.org/10.1016/S0021-9924(03)00019-4

Bansi, S., Oudega, M., Koornneef, A., & van den Broek, P. (2016). *The influence of presentation medium and induced beliefs on reading comprehension: An eyetracking study*. Poster session presented at the Scandinavian Workshop on Applied Eye Tracking 2016, Turku, Finland.

Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press. (Original work published 1932)

Bare, J., & Meek, A. (1998). Issue Brief: Internet Access in U.S. Public Schools (FRSS 69) (NCES 98-031). Washington D.C.: National Center for Education Statistics.

Baron, N. S. (2013). Redefining reading: The impact of digital communication media. *PMLA, 128*, 193-200. https://doi.org/10.1632/pmla.2013.128.1.193

Baron, N. S. (2015). *Words onscreen: The fate of reading in a digital world*. Oxford: Oxford University Press.

Baron, N. S., Calixte, R. M., & Havewala, M. (2017). The persistence of print among university students: An exploratory study. *Telematics and Informatics*, *34*, 590-604. https://doi.org/10.1016/j.tele.2016.11.008

Becker, H. J. (2000). Who's wired and who's not: Children's access to and use of computer technology. *The Future of Children*, *10*, 44-75. https://doi.org/10.2307/1602689

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014). The influence of text annotation tools on print and digital reading comprehension. *Proceedings of the 9th Chais Conference for Innovation in Learning Technologies,* 28-35. http://innovation.openu.ac.il/chais2014/download/B2-2.pdf

Bergen, L., Grimes, T., & Potter, D. (2005). How attention partitions itself during simultaneous message presentations. *Human Communication Research*, *31*, 311-336. https://doi.org/10.1111/j.1468-2958.2005.tb00874.x

Birkerts, S. (1994). *The Gutenberg elegies: the fate of reading in an electronic age*. New York, NY: Fawcett Columbine.

Buzzetto-More, N., Guy, R., & Elobaid, M. (2007). Reading in a digital age: e-books are students ready for this learning object?. *Interdisciplinary Journal of E-Learning and Learning Objects*, *3*, 239-250. https://www.learntechlib.org/p/44809/

Carr, N. (2010). *The shallows: How the internet is changing the way we think, read and remember*. London: Atlantic Books Ltd.

Carrier, L. M., Rosen, L. D., Cheever, N. A., & Lim, A. F. (2015). Causes, effects, and practicalities of everyday multitasking. *Developmental Review, 35*, 64-78. https://doi.org/10.1016/j.dr.2014.12.005

Chen, D., & Catrambone, R. (2015). Paper vs. screen effects on reading comprehension, metacognition, and reader behavior. *Proceedings of The Human Factors and Ergonomics Society Annual Meeting, 59*, 332-336. https://doi.org/10.1177/1541931215591069

Chen, G., Cheng, W., Chang, T., Zheng, X., & Huang, R. (2014). A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education, 1,* 213-225. https://doi.org/10.1007/s40692-014-0012-z

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational*

Research Review, *7*, 198-215. https://doi.org/10.1016/j.edurev.2012.05.002

Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88-113. https://doi.org/10.1016/j.edurev.2013.01.001

Chulkov, D. V., & VanAlstine, J. (2013). College student choice among electronic and printed textbook options. *Journal of Education for Business*, *88*, 216-222. https://doi.org/10.1080/08832323.2012.672936

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*, 445-459. https://doi.org/0.3102/00346543053004445

Clark, R. E. (1994). *Media will never influence learning. Educational Technology Research and Development, 42*(2), 21-29. https://doi.org/10.1007/BF02299088

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, *42*, 214-257. https://doi.org/10.1598/RRQ.42.2.2

Cuban, L., Kirkpatrick, H., & Peck, C. (2001). High access and low use of technologies in high school classrooms: Explaining an apparent paradox. *American Educational Research Journal*, *38*, 813-834. https://doi.org/10.3102/00028312038004813

Daniel, D. B., & Woody, W. D. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education*, *62*, 18-23.

https://doi.org/10.1016/j.compedu.2012.10.016

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, *35*, 1297-1326. https://doi.org/10.1080/00140139208967394

Dobler, E. (2015). e-Textbooks: A Personalized Learning Experience or a Digital Distraction?. *Journal of Adolescent & Adult Literacy*, *58*, 482-491. https://doi.org/10.1002/jaal.391

Eickelmann, B., & Vennemann, M. (2017). Teachers 'attitudes and beliefs regarding ICT in teaching and learning in European countries. *European Educational Research Journal*, *16*, 733-761. https://doi.org/10.1177/1474904117725899

Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, *13*, 93-106. https://www.learntechlib.org/primary/p/4793/

Farinosi, M., Lim, C., & Roll, J. (2016). Book or screen, pen or keyboard? A cross-cultural sociological analysis of writing and reading habits basing on Germany, Italy and the UK. *Telematics and Informatics*, *33*, 410-421. https://doi.org/10.1016/j.tele.2015.09.006

Fiedler, K., Ackerman, R., & Scarampi, C. (2019). Metacognition: Monitoring and controlling one's own knowledge, reasoning and decisions. In R. J. Sternberg & J. Funke (Eds.), *The Psychology of Human Thought: An Introduction* (pp. 89-111). Heidelberg: Heidelberg University Publishing.

Foasberg, N. M. (2014). Student reading practices in print and electronic media. *College & Research Libraries*, *75*, 705-723. https://doi.org/10.5860/crl.75.5.705

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences*, *103*, 11778-11783. https://doi.org/10.1073/pnas.0602659103

Fox, A. B., Rosen, J., & Crawford, M. (2009). Distractions, distractions: Does instant messaging affect college students' performance on a concurrent reading comprehension task?. *CyberPsychology & Behavior, 12*, 51-53. https://doi.org/10.1089/cpb.2008.0107

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age. The IEA International Computer and Information Literacy Study International Report*. Amsterdam: IEA

Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers & Education*, *50*, 906-914. https://doi.org/10.1016/j.compedu.2006.09.006

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3-26). New York, NY: Psychology Press.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395. https://doi.org/10.1037/0033-295X.101.3.371

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163-189. https://doi.org/10.1146/annurev.psych.48.1.163

Graham, M., & Dutton, W. H. (Eds.). (2014). *Society and the internet: How networks of information and communication are changing our lives*. Oxford: Oxford University Press.

Greenfield, P. M. (2009). Technology and informal education: What is taught, what is learned. *Science*, *323*, 69-71. https://doi.org/10.1126/science.1167190

Grimshaw, S., Dungworth, N., McKnight, C., & Morris, A. (2007). Electronic books: Children's reading and comprehension. *British Journal of Educational Technology*, *38*, 583-599. https://doi.org/0.1111/j.1467-8535.2006.00640.x

Guha, S. (2003). Are We All Technically Prepared? Teachers' Perspective on the Causes of Comfort or Discomfort in Using Computers at Elementary Grade Teaching. *Information Technology in Childhood Education Annual, 2003*, 317-349. https://www.learntechlib.org/primary/p/17770/

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, *32*, 139-156. https://doi.org/10.1111/jcal.12123

Haddock, G., Foad, C., Saul, V., Brown, W., & Thompson, R. (2019). The medium can influence the message: Print-based versus digital reading influences how people process different types of written information. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12415

Healy, J. M. (1999). *Failure to connect: How computers affect our children's minds-for better and worse*. New York, NY: Simon & Schuster.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education, 15*, 46-64. https://doi.org/10.1007/BF02940852

Hermena, E. W., Sheen, M., AlJassmi, M., AlFalasi, K., AlMatroushi, M., & Jordan, T. R. (2017). Reading rate and comprehension for text presented on tablet and paper: Evidence from Arabic. *Frontiers in Psychology, 8*, 257. https://doi.org/10.3389/fpsyg.2017.00257

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *The Journal of Technology, Learning, and Assessment, 3* (4). https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1657

Hwang, Y., Kim, H., & Jeong, S. H. (2014). Why do media users multitask?: Motives for general, medium-specific, and content-specific types of multitasking. *Computers in Human Behavior*, *36*, 542-548. https://doi.org/10.1016/j.chb.2014.04.040

Hu, X., Gong, Y., Lai, C., & Leung, F. K. (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Computers & Education*, *125*, 1-13. https://doi.org/10.1016/j.compedu.2018.05.021

Jarvis, J. (2006). Books will disappear. Print is where words go to die. *BBC News*, 15 September 2006. https://www.theguardian.com/commentisfree/2006/jun/05/bookscomment.mediaguardian1

Jeong, H. (2012). A comparison of the influence of electronic books and paper books on reading comprehension, eye fatigue, and perception. *The Electronic Library, 30*, 390-408. https://doi.org/10.1108/02640471211241663

Ji, S. W., Michaels, S., & Waterman, D. (2014). Print vs. electronic readings in college courses: Cost-efficiency and perceived learning. *The Internet and Higher Education*, *21*, 17-24. https://doi.org/10.1016/j.iheduc.2013.10.004

Judge, S., Puckett, K., & Bell, S. M. (2006). Closing the digital divide: Update from the early childhood longitudinal study. *The Journal of Educational Research*, *100*, 52-60. https://doi.org/10.3200/JOER.100.1.52-60

Junco, R., & Cotten, S. R. (2012). No A 4 U: The relationship between multitasking and academic performance. *Computers & Education, 59*, 505-514. https://doi.org/10.1016/j.compedu.2011.12.023

Kendeou, P., McMaster, K. L., & Christ, T. J. (2016). Reading comprehension: Core components and processes. *Policy Insights from the Behavioral and Brain Sciences, 3*, 62-69. https://doi.org/10.1177/2372732215624707

Kingston, N.M. (2008). Comparability of computer- and paper- administered multiple-choice tests for K–12 populations: A synthesis. *Applied Measurement in Education*, *22*, 22-37. https://doi.org/10.1080/08957340802558326

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. *Discourse Processes*, *16*, 193-202. https://doi.org/10.1080/01638539309544837

Kintsch, W (1998). *Comprehension. A paradigm for cognition*. Cambridge: Cambridge University Press.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394. https://doi.org/10.1037/0033-295X.85.5.363

Kirkwood, A., & Price, L. (2014). Technology-enhanced learning and teaching in higher education: what is ‘enhanced’ and how do we know? A critical literature review. *Learning, Media and Technology*, *39*, 6-36.

https://doi.org/10.1080/17439884.2013.770404

Koch, I., Lawo, V., Fels, J., & Vorländer, M. (2011). Switching in the cocktail party: Exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1140-1147. https://doi.org/10.1037/a0022189

Kozma, R. B. (1991). Learning with media. *Review of Educational Research*, *61*, 179-211. https://doi.org/10.3102/00346543061002179

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development, 42*(2), 7-19. https://doi.org/10.1007/BF02299087

Kraushaar, J. M., & Novak, D. C. (2019). Examining the affects of student multitasking with laptops during the lecture. *Journal of Information Systems Education*, *21*, 11. https://aisel.aisnet.org/jise/vol21/iss2/11

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PloS ONE, 8*, e56178. https://doi.org/10.1371/journal.pone.0056178

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. *Journal of Educational Psychology, 75*, 19-26. https://doi.org/10.1037/0022-0663.75.1.19

Kurata, K., Ishita, E., Miyata, Y., & Minami, Y. (2017). Print or digital? Reading behavior and preferences in Japan. *Journal of the Association for Information Science and Technology*, *68*, 884-894. https://doi.org/10.1002/asi.23712

Lanham, R.A. (1995). Digital literacy. *Scientific American, 273*, 198-200. www.jstor.org/stable/24981798

Latini, N., Bråten, I., & Salmerón, L. (2019). It's All about Sex: Reading Medium Affects Processing and Integration of Textual and Pictorial Information about Human Sexuality. Manuscript submitted for publication.

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior*, *35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Liew, C. L., Foo, S., & Chennupati, K. R. (2000). A study of graduate student end-users' use and perception of electronic journals. *Online Information Review*, *24*, 302-315. https://doi.org/10.1108/14684520010350650

Linderholm, T., Virtue, S., Tzeng, Y., & Van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes, 37*, 165-186. https://doi.org/10.1207/s15326950dp3702_5

Lei, J. (2010). Quantity versus quality: A new approach to examine the relationship between technology use and student outcomes. *British Journal of Educational Technology, 41*, 455-472. https://doi.org/10.1111/j.1467-8535.2009.00961.x

Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse Processes, 54*, 427-445. https://doi.org/10.1080/0163853X.2017.1319653

Levine, L. E., Waite, B. M., & Bowman, L. L. (2007). Electronic media use, reading, and academic distractibility in college youth. *Cyberpsychology & Behavior*, *10*, 560-566. https://doi.org/10.1089/cpb.2007.9990

Liang, T., & Huang, Y. (2014). An investigation of reading rate patterns and retrieval outcomes of elementary school students with E-books. *Journal of Educational Technology & Society, 17*, 218. https://search.proquest.com/docview/1502989176

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of documentation, 61, 700-712.* https://doi.org/10.1108/00220410510632040

Liu, Z. (2006). Print vs. electronic resources: A study of user perceptions, preferences, and use. *Information Processing & Management*, *42*, 583-592. https://doi.org/10.1016/j.ipm.2004.12.002

Mangen, A., & Van der Weel, A. (2016). The evolution of reading in the age of digitisation: an integrative framework for reading research. *Literacy, 50*, 116-124. https://doi.org/10.1111/lit.12086

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*, 367-378. https://doi.org/10.1016/S0169-8141(01)00043-9

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology, 35*, 229-241. https://doi.org/10.1016/j.cedpsych.2009.12.001

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100. https://doi.org/10.1006/cogp.1999.0734

Mizrachi, D., Salaz, A. M., Kurbanoglu, S., & Boustany, J., on behalf of the ARFIS Research Group. (2018). Academic reading format preferences and behaviors among university students worldwide: A comparative survey analysis. *PloS ONE*, *13*, e0197444. https://doi.org/10.1371/journal.pone.0197444

Molenda, M. (2008). Historical foundations. In J. M. Spector, M. D. Merrill, J. van Merrienboer, & M. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 3-20). Oxon: Taylor & Francis.

Morris, E. (2019) *Edison*. New York, NY: Random House

Muir, L., & Hawes, G. (2013). The case for e-book literacy: Undergraduate students' experience with e-books for course work. *The Journal of Academic Librarianship*, *39*, 260-274. https://doi.org/10.1016/j.acalib.2013.01.002

Mulet, J., Van de Leemput, C., & Amadieu, F. (2019). A Critical Literature Review of Perceptions of Tablets for Learning in Primary and Secondary Schools. *Educational Psychology Review*, *31*, 631-662. https://doi.org/10.1007/s10648-019-09478-0

Niccoli, A. (2015). Paper or tablet reading recall and comprehension. *EDUCAUSE Review Online, 50.* https://er.educause.edu/articles/2015/9/paper-or-tablet-reading-recall-and-comprehension

Nishizaki, D. M. (2015). *The effects of tablets on learning: Does studying from a tablet computer affect student learning differently across educational levels* (Senior thesis). http://scholarship.claremont.edu/cmc_theses/1011/

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media?. *Computers in Human Behavior*, *54*, 301-309. https://doi.org/10.1016/j.chb.2015.07.043

OECD (2015). *Students, computers and learning: Making the connection*. Paris: OECD.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*, 15583-15587. https://doi.org/10.1073/pnas.0903620106

Pedró, F. (2012) Trusting the unknown: The effects of technology use in Education. In S. Dutta & B. Bilbao (Eds.), *The Global Informatiton Technology Report 2012: Living in a hyperconnected world*. Geneva: World Economic Forum and INSEAD. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.357.1969&rep=rep1&type=pdf

Pernice, K., Whitenton, K., Nielson J. (2014) *How people read on the web: The eyetracting evidence*. Fremont, CA: Nielsen Norman Group.

Ponzo, M. (2011). Does the way in which students use computers affect their school performance? *Journal of Economic & Social Research*, *13*(2), 1-27. https://search.proquest.com/docview/1372334873

Porion, A., Aparicio, X., Megalakaki, O., Robert, A., & Baccino, T. (2016). The impact of paper-based versus computerized presentation on text comprehension and memorization. *Computers in Human Behavior, 54*, 569-576. https://doi.org/10.1016/j.chb.2015.08.002

Potnis, D., Deosthali, K., & Pino, J. (2017). Investigating barriers to "using information" in electronic resources: A study with e-book users. *Proceedings of* *the Association for Information Science and Technology, 54*, 318-326. https://doi.org/10.1002/pra2.2017.14505401035

Potocki, A., Sanchez, M., Ecalle, J., & Magnan, A. (2017). Linguistic and cognitive profiles of 8-to 15-year-old children with specific reading comprehension difficulties: The role of executive functions. *Journal of Learning Disabilities, 50*, 128-142. https://doi.org/10.1177/0022219415613080

Rideout, V. J. (2015). *The common sense census: Media use by tweens and teens*. Common Sense Media Incorporated. https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2015

Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Generation M 2: Media in the lives of 8-to 18-year-olds.* Menlo Park, CA: Henry J. Kaiser Family Foundation.

Rose, E. (2011). The phenomenology of on-screen reading: University students' lived experience of digitised text. *British Journal of Educational Technology*, *42*, 515-526. https://doi.org/10.1111/j.1467-8535.2009.01043.x

Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. *Computers in Human Behavior*, *29*, 948-958. https://doi.org/10.1016/j.chb.2012.12.001

Rouet, J. F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19-52). Charlotte, NC: IAP.

Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, *52*, 200-215. https://doi.org/10.1080/00461520.2017.1329015

Sackstein, S., Spark, L., & Jenkins, A. (2015). Are e-books effective tools for learning? reading speed and comprehension: iPad vs. paper. *South African Journal of Education, 35*(4), 1-14. https://doi.org/10.15700/saje.v35n4a1202

Salmerón, L., Strømsø, H., Kammerer, Y., & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek, (2018). *Learning to read in a digital world*. Amsterdam: John Benjamins Publishing Company.

Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education*, *62*, 24-31. https://doi.org/10.1016/j.compedu.2012.10.003

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS ONE*, *8*, e54402. https://doi.org/10.1371/journal.pone.0054402

Sathe, N. A., Grady, J. L., & Giuse, N. B. (2002). Print versus electronic journals: a preliminary investigation into the effect of journal format on research processes. *Journal of the Medical Library Association*, *90*, 235-243. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC100770/

Seehafer, H. (2014). Effects of learning style on paper versus computer based reading comprehension. *The Red River Psychology Journal, 2014*(1). https://www.mnstate.edu/RRpsychjournal/

Shepperd, J. A., Grace, J. L., & Koch, E. J. (2008). Evaluating the electronic textbook: Is it time to dispense with the paper text?. *Teaching of Psychology*, *35*, 2-5. https://doi.org/10.1080/00986280701818532

Simian, M., Malbrán, C., Fonseca, L., Grinberg, S., Gattas, M. J., Tascón, R., & Cirigliano, S. (2016). *Reading narrative, expository and discontinuous texts on paper versus screen: Impact on reading comprehension in Argentine high school kids*. Poster session presented at 33rd annual meeting of the Society for the Scientific Studies of Reading (SSSR), Porto, Portugal.

Segers, E. (2017). Children's hypertext comprehension. In E. Segers & P. van den Broek, *Developmental perspectives in written literacy: In honor of Ludo Verhoeven*, (pp. 149-164). Amsterdam: John Benjamins Publishing Company

Sidorkin, A. M. (2017). Human capital and innovations in education. In A. M. Sidorkin & M. K. Warford (Eds), *Reforms and Innovation in Education. Implications for the Quality of Human Capitol* (pp. 127-139). Cham: Springer.

Singer, L. M., & Alexander, P. A. (2017a). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research*, *87*, 1007-1041. https://doi.org/10.3102/0034654317722961

Singer, L. M., & Alexander, P. A. (2017b). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education, 85*, 155-172. https://doi.org/10.1080/00220973.2016.1143794

Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language*, *33*, 421-441. https://doi.org/10.1006/jmla.1994.1020

Skryabin, M., Zhang, J., Liu, L., & Zhang, D. (2015). How the ICT development level and usage influence student achievement in reading, mathematics, and science. *Computers & Education*, *85*, 49-58.

https://doi.org/10.1016/j.compedu.2015.02.004

Spector, J. M., Merrill, M. D., Elen, J., & Bishop, M. J. (2014). *Handbook of research on educational communications and technology* (pp. 413-424). New York, NY: Springer.

Spiro, R. J., & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R.J. Spiro, R. J. (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology*. London: LEA.

Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R., Abrami, P. C., Wade, A., et al. (2009). Technology's effect on achievement in higher education: A stage I meta-analysis of classroom applications. *Journal of Computing in Higher Education, 21*, 95–109. https://doi.org/10.1007/s12528-009-9021-8.

Smallwood, J. (2013). Distinguishing how from why the mind wanders: a process–occurrence framework for self-generated mental activity. *Psychological Bulletin, 139*, 519-535. https://doi.org/10.1037/a0030010

Smerdon, B., Cronen, S., Lanahan, L., Anderson, J., Iannotti, N., & Angeles, J. (2000). Teachers' Tools for the 21st Century: A Report on Teachers' Use of Technology. Washington, DC: National Center for Education Statistics.

Snow, C. E. (2002). Reading for understanding: Toward a research and development program in reading comprehension. Santa Monica, CA: RAND.

Sommers, S. R., Shin, L. M., Greenebaum, S. L., Merker, J., & Sanders, A. S. (2019). Quasi-experimental and experimental assessment of electronic textbook experiences: Student perceptions and test performance. *Scholarship of Teaching and Learning in Psychology*, *5*, 11-22. https://doi.org/10.1037/stl0000129

Stein, N. L., & Trabasso, T. (1985). The search after meaning: Comprehension and comprehension monitoring. In F. J. Morrison, C. Lord, & D. Keating (Eds.), *Applied developmental psychology* (Vol 2., pp. 33-58). San Diego, CA: Academic Press.

Stoop, J., Kreutzer, P., & Kircz, J. (2013). Reading and learning from screens versus print: a study in changing habits: Part 1–reading long information rich texts. *New Library World*, *114*, 284-300. https://doi.org/10.1108/NLW-04-2013-0034

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research*, *81*, 4-28. https://doi.org/10.3102/0034654310393361

Taylor, A. K. (2011). Students learn equally well from digital as from paperbound texts. *Teaching of Psychology, 38*, 278-281. https://doi.org/10.1177/0098628311421330

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*, 453-456. https://doi.org/10.1126/science.1182594

van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. Magliano, & G. Schraw (Eds.), *Relevance instructions and goal- focusing in text learning* (pp. 123-140). Greenwich, CT: Information Age Publishing.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*, 360-372.

https://doi.org/10.1080/0163853X.2017.1306677

van den Broek, P., Helder, A., & Van Leijenhorst, L. (2013). Sensitivity to structural centrality: Developmental and individual differences in reading comprehension skills. In M. A. Britt, S. R. Goldman, & J. F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. 132-146). New York, NY: Routledge.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*, 335-351. https://doi.org/10.1002/acp.1418

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, *39*, 299-316. https://doi.org/10.1080/0163853X.2005.9651685

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp & S.R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Erlbaum.

van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior, 53*, 204-215. https://doi.org/10.1016/j.chb.2015.06.035

van der Weel, A. (2011). *Changing our textual minds. Towards a digital order of knowledge.* Manchester: Manchester University Press.

Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). Strategies of discourse comprehension. New York, NY: Academic Press.

Van Horne, S., Russell, J. E., & Schuh, K. L. (2016). The adoption of mark-up tools in an interactive e-textbook reader. *Educational Technology Research and Development*, *64*, 407-433. https://doi.org/10.1007/s11423-016-9425-x

Wang, Y. M., & Holthaus, P. (1999). Facing the world: Student teachers' computer use during practicum. *Journal of Educational Technology Systems*, *27*, 207-223. https://doi.org/10.2190/HB1N-0HF0-QFCM-2CVV

Wang, S., Jiao, H., Young, M.J., Brooks, T., & Olson, O. (2007). Comparability of computer-based and paper-and-pencil testing in K–12 reading assessments. A meta-analysis of testing mode effects. *Educational and Psychological Measurement, 68*, 5-24. https://doi.org/10.1177/0013164407305592

Wenglinsky, H. (2005). *Using technology wisely: The keys to success in schools*. New York, NY: Teachers College Press.

Wharton, C., & Kintsch, W. (1991). An overview of construction-integration model: a theory of comprehension as a foundation for a new cognitive architecture. *ACM SIGART Bulletin*, *2*(4), 169-173. https://doi.org/10.1145/122344.122379

Wolf, M. (2018). *Reader, come home: The reading brain in a digital world*. New York, NY: HarperCollins

Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers & Education*, *58*, 365-374. https://doi.org/10.1016/j.compedu.2011.08.029

Woody, W. D., Daniel, D. B., & Baker, C. A. (2010). E-books or textbooks: Students prefer textbooks. *Computers & Education*, *55*, 945-948. https://doi.org/10.1016/j.compedu.2010.04.005

Wylie, J., Thomson, J., Leppänen, P. H., Ackerman, R., Kanniainen, L., & Prieler, T. (2018). Cognitive processes and digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek, (2018). *Learning to read in a digital world*. Amsterdam: John Benjamins Publishing Company.

Chapter 2

Critical analysis of the effects of the digital technologies on reading and learning*

Ladislao Salmerón and Pablo Delgado

~

"The problem with the technology-centered approach is that it is more concerned with promoting educational technology than with promoting learning in students."

Richard Mayer, *An interview with Richard Mayer* (Veronikas & Shaughnessy, 2005)

* The content of this chapter has been published under the reference:

Salmerón, L., & Delgado, P. (2019). Critical analysis of the effects of the digital technologies on reading and learning / Análisis crítico sobre los efectos de las tecnologías digitales en la lectura y el aprendizaje. *Cultura y Educación*, *31*, 465-480. https://doi.org/10.1080/11356405.2019.1630958

Chapter 2

Critical analysis of the effects of the digital technologies on reading and learning*

IN THE DIGITAL AGE, the introduction of the digital technologies (or information and communication technologies, ICT) in classrooms is a necessary, imperative reality. However, the educational results do not seem to match the expectations generated. For this reason, this revision seeks to make up for the scant attention paid to the possible harmful effects of the use of ICT based on research and educational practice. After providing a general overview of the impact of ICT on learning, we survey the results of studies that reveal potential negative consequences of their use inside and outside the school context. We analyse their impact on both academic performance and other more specific areas, such as reading comprehension, and we stress the negative effects on so-called "digital natives". In short, what stands out is the importance of an evidence-based education practice which bears in mind the possible harmful effects of using ICT and bears in mind that its beneficial effects seem to rely on not only how much but also how they are used.

* This chapter was translated to English by Mary Black

2.1. Introduction

In the past three decades, the relationship between the digital technologies and school education has evolved from an initial orientation which sought to teach students how to use digital technology in the 1980s and 1990s to a vision of permeation, which has been dominant in the 21st century, that tries to get students to learn more and better through the use of the information and communication technologies (ICT) (Vivancos, 2008). However, in addition to a widespread tendency to include ICT without assessing their efficacy (Sidorkin, 2017), the stress on its potential benefits (e.g., Archer et al., 2014; Cheung & Slavin, 2013) contrasts with the lack of critical visions on its possible negative effects (Goodchild & Speed, 2018). Yet we should not ignore the fact that the ultimate goal of any educational innovation is for students to learn better. In this regard, not only are there numerous scholarly studies that indicate that merely including ICT does not improve learning, but the increasing use of digital devices may have negative repercussions, as we shall discuss in more depth throughout this article. Specifically, the sections below analyse the studies that have identified negative relationships between the digital technologies, reading and learning. First, we survey the relevant literature on the use of digital technology in the classroom. Secondly, we discuss the role of the social media in learning. Thirdly, we analyse the role of multitasking and digital technology in learning. And finally, we conclude with recommendations on the use of the digital technologies in education.

2.2. Digital technologies and their use at school

Andreas Schleicher, the Education Division Head of the OECD and the person in charge of the International Student Assessment (PISA) programme, recently stated that digital technology has aroused too many false hopes (Schleicher, 2015). The PISA results speak volumes: not only are the learning results not better, but in some cases the use of ICT is associated with poor performance. These statements are based on the evaluation of approximately half a million students aged 15 to 16 who are representative of the student body in each of the 65 participating countries (OECD, 2015). Therefore, these are not occasional cases of success or failure in the implementation of digital technology in the classroom but patterns which can be generalised to the majority of OECD countries.

The digital reading competence has been defined as the ability to locate, integrate and reflect on information in a digital format (OECD, 2009; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018). Even though we could expect reading practice in digital environments to foster the acquisition and improvement of this skill, the PISA results from 2012 (OECD, 2015) indicate that students who often use computers at school (more than 'once or twice a week') show a lower digital reading competence than those who do not use them or only use them moderately. This same negative pattern was found in the results of the 2012 PISA test (OECD, 2015, chapter 6) in relation to mathematical and scientific competences, defined as specific knowledge and skills to apply the contents of these fields to problems in everyday life (OECD, 2009). Specifically, the students who often use computers in class tend to score lower on mathematical and scientific competence tests than those who either do not use them or use them infrequently.

Three important conclusions can be drawn from these results: 1) mere practice with computers does not guarantee the development of the digital reading competence, which implies the need to analyse how technology is used, more than how much; and more worrisome, 2) a high use of computers at school seems to be associated with poor performance; and 3) this does not seem to be an isolated phenomenon, since it has been reported in the fields of reading, mathematics and science. Therefore, this is a complex phenomenon. Given that we are not referring to an addictive use of digital technologies but to a supposedly academic practice, it is plausible to interpret that this use is not being planned with effective methodologies. For example, the results of an international survey conducted in a representative sample of teachers and students in grades equivalent to the second year of secondary school in 21 countries indicate that ICT is primarily used for simple tasks by both students (writing essays and class presentations) and teachers (supporting their explanations and reinforcing examples through repetition) (Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014).

Given this situation, we should ask why digital technology continues to seem so attractive at school, and why is its mass inclusion considered imperative? It is unquestionable that its inclusion in the classroom originally reflected a social need. In a society where the presence of this technology is ubiquitous, schools should educate technologically competent students. However, the fact that students have to learn from and with digital technologies does not mean that their omnipresence in school should be accepted, especially when their effects may be pernicious. Therefore, to understand the mass inclusion of digital technology in the classroom, we should look at other types of factors. In our viewpoint, both technology lobbies (as Goodchild and Speed (2018) have recently criticised) and the power of suggestion of the digital technologies themselves may play a prominent role.

One intriguing example of digital technologies is the Promethean 2008-2010 study on the use of interactive digital boards (IDB) (Coscollola, 2011), in which 85 teachers and 3,400 primary and secondary school students participated. At the end of the first academic year in which IDBs were used in the classrooms, 90% of the teachers believed that the IDB: a) increased students' attention, motivation and participation; b) facilitated understanding; and c) enhanced skills for making presentations, argumentations and corrections. Furthermore, and more surprising, even though the academic results were very similar to those from the previous year, the teachers erroneously believed that the students had learned more. This mismatch is even more worrisome if we consider that teaching professionals tend to base the adoption of digital technologies on the experiences of other colleagues instead of on recommendations based on studies or reports like PISA (Price & Kirkwood, 2014).

On the other hand, economic pressure can be exemplified by apparently scientific studies, like the "Tablets in Education" report (Camacho, 2017). Financed by a technology multinational, it summarises the experience of implementing digital tablets in Spanish schools between 2014 and 2017 with the goal of studying their impact on several competences, such as language and digital skills. Unfortunately, the study does not meet the parameters expected of a relevant scholarly study: the measures used to evaluate competences are not described, no control group is used, and the results are not statistically analysed. As such, it seems more like a marketing product than an educational report, which did not stop it from being presented at the National Library or prevent the Ministry of Education from broadly disseminating it (INTEF, 2017).

Therefore, we believe it is essential for the educational community to get critically involved in analysing the effects of digital technology in the classroom while avoiding supposedly foregone conclusions without the support of evidence (de Pablos, 2008). This is no easy task when a biased vision is promoted even by the institutions themselves. In short, it does not seem essential to inundate schools with technology to develop the digital competence, and indeed an intensive use of technology can be harmful. Does the same hold true of other uses of the digital technologies, such as the social media?

2.3. The social media, reading and learning

The effects of the digital technologies on reading and learning extend beyond school. It is enlightening to distinguish between two main uses of digital technology in general and the Internet in particular: social use and informative use (Naumann, 2015). While the former entails the use of the social media, chats, informal emails and online games, the informative use refers to information searches on websites, blogs or forums to learn. As we shall see, each is associated with very different effects on learning.

Thus, the use of the social media appears to be directly related to lower academic performance, as shown in a recent meta-analysis carried out with data on 21,367 students (Huang, 2018). Specifically, the amount of time that students spend daily using the social media is negatively associated with their academic performance, as measured by their marks. This same negative pattern was also found in relation to their digital reading competence (Borgonovi, 2016; Lee & Wu, 2013; Pfost, Hattie, Dörfler, & Artelt, 2014). Why does using the social media lead to worse student performance? There are two possible explanations. On the one hand, the amount of time the students spend in those environments should be considered a key factor more than how many times they use them (Huang, 2018). This suggests that the amount of time on the social media limits the amount of time spent studying, which would explain the lower academic performance. On the other hand, the use of the social media tends to be associated with a constant and immediate flow of superficial information. Backing this are studies that find that using the social media frequently is related to less reflective thinking, as measured through selfreporting (Annisette & Lafreniere, 2017).

With regard to the negative impact on the digital reading competence, it could be related to the type of information typically accessed on the social media and the language in which it is written, which is more closely associated with oral language, which is not as linguistically and cognitively demanding as what is needed to understand academic language (Snow, 2010). In contrast, it has been reported that the informational use of digital technology, such as reading articles from Wikipedia when doing homework, is indeed associated with the development of this digital competence (Naumann, 2015; Salmerón, García, & Vidal-Abarca, 2018).

A different problem emerges when the students have access to the social media in the classroom, which can lead to distractions that interfere with the learning process. In an experiment with university students, Demirbilek and Talan (2017) allowed one group to use the social media during a class, specifically to use Facebook and to send brief messages to their classmates. The control group was not allowed access to the social media. The students' learning, as measured by a test on the contents of the class, was higher in the group without access to the social media. Other qualitative studies provide further nuances on this phenomenon. Specifically, Karpinski, Kirschner, Ozer, Mellott and Ochwo (2013) found that the negative effect of using the social media during class depends on how they are used. The students who used them for topics unrelated to the class learned less than those who did not use them, while those who used them for academic purposes learned at the same level as those who did not use them. The non-academic use of the social media in class negatively affects the majority of students, even those with high intelligence, motivation or interest in the subject (Ravizza, Hambrick, & Fenn, 2014; Ravizza, Uitvlugt, & Fenn, 2017).

In short, we cannot trust that the way students are using the digital technologies in social media prepares them for better classroom learning. Furthermore, the use of the social media points to an additional problem, as their simultaneous use with learning processes may interfere with academic performance. This phenomenon may be considered a kind of multitasking, which has been studied in depth in the past decade when attempting to describe the characteristics of digital natives.

2.4. The digital technologies and multitasking

The notion of 'digital natives' suggests that those who have grown up surrounded by digital technology possess different, advanced information-processing capacities, including a special skill for engaging in different tasks concurrently (Prensky, 2001; Veen & Vrakking, 2006). This vision, though quite widespread, has been systematically refuted by research. Most youths today are not skilled at handling digital information (Fajardo, Villalta, & Salmerón, 2016; Kirschner & De Bruyckere, 2017), nor are they capable of successfully multitasking, given that the human brain cannot engage in two simultaneous tasks without diminishing performance on one of them (Dindar & Akbulut, 2016). In fact, brain activity does not seem to change qualitatively to adapt to multitasking, but instead there is a greater activation of the left prefrontal cortex, which is involved in controlling sustained attention (Moisala et al., 2016). The brain of adolescents who multitask must make a greater effort, and not always successfully, to prevent the associated effects of distraction.

Nonetheless, digital multitasking is quite widespread amongst youths (Carrier, Cheever, Rosen, Benitez, & Chang, 2009; Rideout, Foehr, & Roberts, 2010; Voorveld & van der Groot, 2013), which has led to an increasing number of studies on its effects on learning. Findings reveal that it is associated with higher distraction in adolescents in activities that require high levels of concentration (Loh, Tan, & Lim, 2016; Moisala et al., 2016). This distracting effect even occurs among university students listening to instructors' lectures when they can see their classmates using laptop computers to do other activities (Sana, Weston, & Cepeda, 2013). Therefore, digital multitasking does not seem recommendable when the purpose is learning; instead, it is preferable to teach students to sequence their tasks and focus on each of them until they achieve their purpose.

What is more, even digital natives understand worse when they read on digital devices (e.g., Lenhard, Schroeders, & Lenhard, 2017; Singer Trakhman, Alexander, & Berkowitz, 2017). Indeed, a recent meta-analysis revealed that reading printed texts leads to better comprehension, especially when there is limited time to read. Furthermore, digital reading's inferiority seems to have grown in the past 17 years (see Study 1 in Chapter 3); in other words, digital natives understand better on paper than on a screen, even compared to previous generations.

The explanation of digital media's inferiority may at least partly come from the existence of certain cognitive self-regulation difficulties when using this medium (e.g., Ackerman & Lauterman, 2012).

2.5. Conclusions

The promise that ICT promote the development of more and better competences is questionable at best. The evidence shows that in many cases using these technologies, as well as many of their associated uses, has zero or even negative effects on the development of the digital reading competence and learning. Knowing these effects is a necessary step towards taking informed decisions on the use of digital technology in the classroom.

This pessimistic view of digital technology should not be understood as a call to disconnect classrooms, a clearly unrealistic proposal in the digital age. Nor can we ignore the studies that cite some of its benefits (e.g., Archer et al., 2014; Cheung & Slavin, 2013). Our goal is simply to warn about the potential risks and to highlight that the introduction of this technology should come hand in hand with evidence-based practices which take advantage of its benefits and instruct students so that they perform successfully in the digital world. Both students and teachers should be aware that using digital technology may be associated with difficulties maintaining attention and acting reflectively. In education, we should promote the development of metacognitive skills so that students can effectively regulate their learning process in the digital environment. For example, Salmerón and Llorens (2018) have used video modelling to teach self-regulated information searches on Wikipedia with secondary school students. The instruction was carried out in pairs, as this is a pedagogical activity that fosters the transfer of learning, which enables students to apply what they have learned in other contexts (Salmerón, 2013). After the instruction, the students went from doing digital speed reading (that is, merely scanning the information) to reading the contents of the websites more carefully, which enabled them to better understand their content.

Likewise, it is important to sequence the activities and avoid multitasking to the extent possible. For example, there are editable digital activities like WebQuests (Argelagós & Pifarré, 2012) which foster self-regulation of the digital competence through guided tasks, yet they also offer some degree of flexibility to promote student autonomy.

Future research should continue to analyse the effects of digital technology on other educational competences, which may also be harmed by the new technology. For example, even though there is not enough evidence to consider it a robust phenomenon, recent studies suggest that students learn more when they write or take notes on paper than when they do so on computers (Alves et al., 2015; Mueller & Oppenheimer, 2014).

In short, even though the relationship between digital technology, reading and learning is not bereft of risks, it is in the hands of the educational community to foster the effective use of ICT to maximise their potentialities and limit the harm they can do.

References

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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Alves, R. A., Limpo, T., Fidalgo, R., Carvalhais, L., Pereira, L. Á., & Castro, S. L. (2016). The impact of promoting transcription on early text production: Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology*, *108*, 665-679. https://doi.org/10.1037/edu0000089

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115,* 54-158. https://doi.org/10.1016/j.paid.2016.02.043

Archer, K., Savage, R., Sanghera-Sidhu, S., Wood, E., Gottardo, A., & Chen, V. (2014). Examining the effectiveness of technology use in classrooms: A tertiary meta-analysis. *Computers & Education, 78*, 140-149. https://doi.org/10.1016/j.compedu.2014.06.001

Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior, 28*, 515-526. https://doi.org/10.1016/j.chb.2011.10.024

Borgonovi, F. (2016). Video gaming and gender differences in digital and printed reading performance among 15-year-olds students in 26 countries. *Journal of Adolescence*, *48*, 45-61. https://doi.org/10.1016/j.adolescence.2016.01.004

Camacho, M. (2017). *Tablets en educación. Hacia un aprendizaje basado en* *competencias*. http://blog.educalab.es/intef/2017/10/04/estudio-tabletas-en-educacion-hacia-un-aprendizaje-basado-en-competencias/

Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of Americans. *Computers in Human Behavior, 25,* 483-489. https://doi.org/10.1016/j.chb.2008.10.012

Coscollola, M. D. (2011). Pizarra Digital Interactiva en el aula: Uso y valoraciones sobre el aprendizaje. *Estudios sobre Educación, 20,* 99-116. http://hdl.handle.net/10171/18345

de Pablos, J. (2008). Algunas reflexiones sobre las tecnologías digitales y su impacto social y educativo. *Quaderns Digitals, 51*.

Demirbilek, M., & Talan, T. (2017). The effect of social media multitasking on classroom performance. *Active Learning in Higher Education, 19*, 117-129. https://doi.org/10.1177/1469787417721382

Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88-113. https://doi.org/10.1016/j.edurev.2013.01.001

Dindar, M., & Akbulut, Y. (2016). Effects of multitasking on retention and topic interest. *Learning and Instruction, 41,* 94-105. https://doi.org/10.1016/j.learninstruc.2015.10.005

Fajardo, I., Villalta, E., & Salmerón, L. (2016). ¿Son realmente tan buenos los nativos digitales? Relación entre las habilidades digitales y la lectura digital. *Anales de Psicología, 32,* 89-97. http://dx.doi.org/10.6018/analesps.32.1.185571

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age. The IEA International Computer and Information Literacy Study International report*. Amsterdam: IEA.

Goodchild, T., & Speed, E. (2018). Technology enhanced learning as transformative innovation: a note on the enduring myth of TEL. *Teaching in Higher Education, 24,* 948-963. https://doi.org/10.1080/13562517.2018.1518900

Huang, C. (2018). Social network site use and academic achievement: A meta-analysis. *Computers & Education, 119,* 76-83. https://doi.org/10.1016/j.compedu.2017.12.010

INTEF (2017). *Estudio "Tablets en educación. Hacia un aprendizaje basado en competencias"*. http://blog.educalab.es/intef/2017/10/04/estudio-tabletas-en-educacion-hacia-un-aprendizaje-basado-en-competencias/

Karpinski, A. C., Kirschner, P. A., Ozer, I., Mellott, J. A., & Ochwo, P. (2013). An exploration of social networking site use, multitasking, and academic performance among United States and European university students. *Computers in Human Behavior, 29,* 1182-1192. https://doi.org/10.1016/j.chb.2012.10.011

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, *67*, 135-142. https://doi.org/10.1016/j.tate.2017.06.001

Lee, Y. H., & Wu, J. Y. (2013). The indirect effects of online social entertainment and information seeking activities on reading literacy. *Computers & Education, 67*, 168-177. https://doi.org/10.1016/j.compedu.2013.03.001

Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse*

Processes, 54, 427-445. https://doi.org/10.1080/0163853X.2017.1319653

Loh, K. K., Tan, B. Z. H., & Lim, S. W. H. (2016). Media multitasking predicts video-recorded lecture learning performance through mind wandering tendencies. *Computers in Human Behavior, 63,* 943-947. https://doi.org/10.1016/j.chb.2016.06.030

Moisala, M., Salmela, V., Hietajärvi, L., Salo, E., Carlson, S., Salonen, O., Lonkab, K., Hakkarainena, K., Salmela-Arog, K., & Alho, K. (2016). Media multitasking is associated with distractibility and increased prefrontal activity in adolescents and young adults. *NeuroImage, 134,* 113-121. https://doi.org/10.1016/j.neuroimage.2016.04.011

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, *25*, 1159-1168. https://doi.org/10.1177/0956797614524581

Naumann, J. (2015). A model of online reading engagement: linking engagement, navigation, and performance in digital reading. *Computers in Human Behavior, 53,* 263-277. https://doi.org/10.1016/j.chb.2015.06.051

OECD (2009). *PISA 2009. Assessment framework: Key competencies in reading, mathematics, and science.* Paris: OECD.

OECD (2015). *Students, computers and learning: Making the connection*. Paris: OECD.

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research, 84,* 203-244. https://doi.org/10.3102/0034654313509492

Prensky, M. (2001). Digital natives, digital immigrants, part 1. *On the Horizon*, 9(5), 1-

6. https://doi.org/10.1108/10748120110424816

Price, L., & Kirkwood, A. (2014). Using technology for teaching and learning in higher education: a critical review of the role of evidence in informing practice. *Higher Education Research and Development, 33,* 549-564. https://doi.org/10.1080/07294360.2013.841643

Ravizza, S. M., Hambrick, D. Z., & Fenn, K. M. (2014). Non-academic internet use in the classroom is negatively related to classroom learning regardless of intellectual ability. *Computers & Education, 78,* 109-114. https://doi.org/10.1016/j.compedu.2014.05.007

Ravizza, S. M., Uitvlugt, M. G., & Fenn, K. M. (2017). Logged in and zoned out: How laptop internet use relates to classroom learning. *Psychological Science, 28,* 171-180. https://doi.org/10.1177/0956797616677314

Rideout, V.J., Foehr, U.G., & Roberts, D.F. (2010). *Generation M2: Media in the Lives of 8- to 18-year-olds.* Menlo Park: The Henry J. Kaiser Family Foundation.

Salmerón, L. (2013). Actividades que promueven la transferencia de los aprendizajes: una revisión de la literatura. *Revista de Educación, volumen extraordinario 2013*, 34-53.

Salmerón, L., García, A., & Vidal-Abarca, E. (2018). The development of adolescents' comprehension-based Internet reading skills. *Learning and Individual Differences, 61,* 31-39. http://roderic.uv.es/handle/10550/34528

Salmerón, L., & Llorens, A. (2018). Instruction of digital reading strategies based on eye movements modeling examples. *Journal of Educational Computing Research, 57,* 343-359. https://doi.org/10.1177/0735633117751605

Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, 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: John Benjamins.

Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24-31. https://doi.org/10.1016/j.compedu.2012.10.003

Sidorkin, M. A. (2017). Human capital and innovations in education. In A. M. Sidorkin y M. K. Wardford (Eds.). *Reforms and innovation in Education. Implications for the quality of human capital* (pp. 127-129). Cham: Springer.

Singer Trakhman, L. M., Alexander, P. A., & Berkowitz, L. E. (2017). Effects of processing time on comprehension and calibration in print and digital mediums. *The Journal of Experimental Education*. https://doi.org/10.1080/00220973.2017.1411877

Schleicher, A. (2015). School technology struggles to make an impact. *BBC News*, 15 September 2015. http://www.bbc.com/news/business-34174795

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328,* 450-452. https://doi.org/10.1126/science.1182597

Veen, W., & Vrakking, B. (2006). *Homo zappiens: Growing up in a digital age*. London: Network Continuum Education.

Vivancos, J. (2008). *Tratamiento de la información y competencia digital.* Madrid: Alianza Editorial.

Voorveld, H.A.M., & van der Groot, M. (2013). Age differences in media multitasking: a diary study. *Journal of Broadcasting & Electronic Media, 57,* 392-408. https://doi.org/10.1080/08838151.2013.816709

Chapter 3

Study 1

Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension*

Pablo Delgado, Cristina Vargas, Rakefet Ackerman, and Ladislao Salmerón

~

"Both the man of science and the man of art live always at the edge of mystery, surrounded by it; both always, as to the measure of their creation, have had to do with the harmonization of what is new with what is familiar, with the balance between novelty and synthesis, with the struggle to make partial order in total chaos."

J. Robert Oppenheimer, *Prospects in the art and sciences* (1954)

* The content of this chapter has been published under the reference:

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. https://doi.org/10.1016/j.edurev.2018.09.003

Chapter 3

Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension

WITH THE INCREASING dominance of digital reading over paper reading, gaining understanding of the effects of the medium on reading comprehension has become critical. However, results from research comparing learning outcomes across printed and digital media are mixed, making conclusions difficult to reach. In the current meta-analysis, we examined research in recent years (2000–2017), comparing the reading of comparable texts on paper and on digital devices. We included studies with between-participants (n = 38) and within-participants designs (n = 16) involving 171,055 participants. Both designs yielded the same advantage of paper over digital reading (Hedge's $g = -0.21$; $d_C = -0.21$). Analyses revealed three significant moderators: (1) time frame: the paper-based reading advantage increased in time-constrained reading compared to self-paced reading; (2) text genre: the paper-based reading advantage was consistent across studies using informational texts, or a mix of informational and narrative texts, but not on those using only narrative texts; (3) publication year: the advantage of paper-based reading increased over the years. Theoretical and educational implications are discussed.

3.1. Introduction

There has been a gradual shift from paper-based reading to reading on digital devices, such as computers, tablets, and cell-phones. Although there are clear advantages of digital-based assessment and learning, including reduced costs and increased individualization, research indicates that there may be disadvantages as well, as described below. In addition, findings from previous reviews of studies on the effects of digital reading on comprehension have been inconclusive (Dillon, 1992; Kingston, 2008; Noyes & Garland, 2008; Singer & Alexander, 2017b; Wang, Jiao, Young, Brooks, & Olson, 2007). The current paper presents a meta-analysis of recent studies that investigated the effects of paper versus digital media on reading comprehension. In addition, we also explored the effects of several potential moderator variables whose influence may help to explain previous inconsistencies among study results.

3.1.1. Text comprehension and the role of media

Theoretical models of reading comprehension have extensively considered the interplay among reader characteristics, text content and design, and reading instructions (for a review see McNamara & Magliano, 2009). However, the factor of the medium has been mostly ignored, despite empirical evidence suggesting that it influences reading outcomes (e.g., Lenhard, Schroeders, & Lenhard, 2017; Mangen, Walgermo, & Brønnick, 2013; Singer & Alexander, 2017a). In particular, Ackerman and Lauterman (2012) considered media-related differences in learning outcomes from a metacognitive perspective. In addition to learning outcomes, they compared learners' monitoring of their comprehension and allocation of their study time. On each medium, immediately after studying each text, participants predicted their success rates (in %) and were tested through multiple-choice questions. Moreover, to the best of our knowledge, these authors are the only ones who empirically considered the time frame as a potential moderating factor of media effects on learning outcomes. They examined the learners' adjustment to studying under time pressure, compared to free study time, on both media. Under time pressure, but not under free time, those who read from computers showed screen inferiority: they had more pronounced overconfidence than paper learners and achieved lower test scores. Moreover, only in paper- based reading, participants improved their efficiency under time pressure, compared to learning in a free time frame. Importantly, whereas theories of monitoring and allocation of study time assume close relationships between the two, Ackerman and Goldsmith (2011) found close relationships in paper-based reading, but more erratic time allocation decisions in digital-based reading. Before this study, conducted with young undergraduates, weak associations between monitoring and time allocation decisions were only found in elderly people and people with mental illnesses (Koren, Sneidman, Goldsmith, & Harvey, 2006; Pansky, Goldsmith, Koriat, & Pearlman-Avnion, 2009). Furthermore, several recent studies found that the preference for paper over digital-based reading persists despite technological advances (Baron, Calixte, & Havewala, 2017; Mizrachi, 2015; Kurata, Ishita, Miyata, & Minami, 2017; but see; Singer & Alexander, 2017a). Lauterman and Ackerman (2014) found that methods to overcome screen inferiority are effective only for people who prefer digital reading, but not for those who prefer paper reading. Together, the reviewed findings demonstrate several aspects of reading comprehension that have been overlooked so far in reading theories, highlighting the medium as an environment that affects reading outcomes, above and beyond reader and task characteristics.

In sum, the way the media affect reading comprehension outcomes is still unclear.

Several researchers have explained screen inferiority under some conditions as being due to people's stronger inclination toward shallow work in digital-based environments than in paper-based ones (see Annisette & Lafreniere, 2017; Wolf & Barzillai, 2009), particularly when the task design indicates its legitimacy, as when working under a limited time frame (Lauterman & Ackerman, 2014; Sidi, Shpigelman, Zalmanov, & Ackerman, 2017).

A meta-analysis provides an opportunity to examine media effects on learning outcomes while considering overall task characteristics, such as time frames, participant characteristics, and the display technology, across theoretical frameworks, populations, and methodologies. Importantly, a meta-analysis makes it possible to consider potentially moderating factors, even across studies that did not include these factors in their designs, by comparing enough studies that used each level of the factor (e.g., only limited time frame vs. only free time allocation). Exposing moderating factors can guide future theoretical development and practical recommendations.

3.1.2. Previous reviews and meta-analyses

In the past ten years, only a few meta-analyses and literature reviews have been undertaken to determine the nature of the medium's influence on reading outcomes. Wang et al. (2007) focused on K-12 student population. Their meta-analysis examined media effects on performance on standardized tests, and it included 11 primary studies that yielded 42 comparisons. They found better reading outcomes in paper-based testing than in digital-based testing. The mean effect size (0.08) was significant, but small (see Cohen, 1988), and this difference between reading media was larger in studies that used fixed linear computerized tests (n = 37) than in those that used adaptive computerized tests (n

= 5). Wang et al. concluded that differences between testing media are probably test specific, so that an analysis of potential media effects should be conducted for each type of test separately.

Kingston (2008) conducted a larger meta-analysis that included 81 effect sizes from 16 studies. This study focused on testing academic achievement across several academic topics in K-12 populations, and it showed a small advantage for digital administration in English Language Arts and Social Studies (effect sizes of 0.11 and 0.15, respectively), along with a small advantage for paper administration in Mathematics (effect size of −0.06). More relevant to our focus, eight of the studies included in Kingston's work assessed reading outcomes, five of which were included in Wang et al.'s (2007) meta-analysis, and found no effect of reading media. Regarding the digital disadvantage in Mathematics, Kingston alludes to possible difficulties when completing tests on a computer due to switching to sketch paper before answering. In sum, results from these meta-analyses are inconsistent. Some findings point to advantages of print text, whereas others favour digital text, and still other results indicate that media effects depend on the topic.

Recently, Kong, Seo, and Zhai (2018) performed a meta-analysis with 17 studies dating from 2000 to 2016. Results revealed better performance when reading from paper than when reading from digital devices (effect size of −0.21). This meta-analysis incorporated a relatively small number of studies which included great variability in terms of populations (e.g. second-language students), and tasks (e.g. perceived comprehension or proofreading). Interestingly, despite considering several potential moderating factors, this analysis did not reveal any significant effects. The authors acknowledged the need for considering additional moderating factors.

Two narrative literature reviews attempted to promote understanding of media effects on reading comprehension. Noyes and Garland (2008) reviewed media comparison studies that focused on reading outcomes but also on tasks such as examinations, writing, and filling in questionnaires (e.g., psychometric tests and surveys). They concluded that, although equivalence between the media was a challenge, differences, where found, appeared to be task specific. In particular, with respect to reading outcomes, the results were heterogeneous regarding comprehension and reading speed, with no clear conclusions about the influence of the media.

Recently, Singer and Alexander (2017b) described studies published from 1992 to 2017. They found it difficult to reach conclusions and pointed to a lack of clarity in definitions of paper and digital reading, as well as a lack of important information in many studies, such as text features (genre and length), individual differences (e.g., reading rate and vocabulary), validity and re- liability of the tasks used to measure reading outcomes, characteristics of the reading tasks, levels of comprehension evaluated, and scoring criteria. Singer and Alexander called on researchers to investigate how various factors interact with media and potentially explain the mixed results found in the literature.

The main conclusion drawn from the above review of previous meta-analyses and narrative research synthesis is that media effects are inconsistent. This may be partially explained by the difficulty of comparing paper texts to digital texts which include incomparable features such as hyperlinks, animations, or adaptive tests which may confound and hide media effects on learning processes. Another potential reason for the inconsistent results is the fact that most of the previous reviews did not consider or did not find moderating factors. Finding robust moderating factors can shed light on the reasons for the seemingly inconsistent media effects found. As mentioned above, Ackerman and Lauterman (2012) found inferior comprehension in digital-based reading compared to paper-based reading under time pressure, but media equivalence in free time conditions. This finding raises the option that the time frame allowed for reading is a factor that differentiates between studies that find an advantage of paper and those that find media equivalence. Considering the time frame as a moderating factor across a large collection of studies can inform us whether this specific study exposed a pattern which is robust across methodologies and populations.

In the present meta-analysis, we aimed to facilitate comparisons between print and digital media by including only studies that used linear reading materials, where the digital texts closely resembled the printed versions. This focus allowed us to eliminate some of the aforementioned complexities. In addition, by performing a comprehensive meta-analysis we aimed to examine the influence of several potential moderating factors on media effects, in addition to the time frame just mentioned. We see high importance in identifying moderating factors for pointing to conditions that yield an advantage of print across methodologies and conditions, those that yield an advantage of digital devices, and those that result in equivalent outcomes.

3.1.3. Effects of experience with digital technologies

It could be argued that a potential straightforward moderator of digital text comprehension is experience using technology. In other words, potential comprehension difficulties in digital reading will disappear once students have enough experience with digital technologies. According to this view, as each new generation is surrounded by digital devices earlier and earlier in life (e.g. ASHA, 2015; Childwise, 2017), we should expect newer generations to achieve equivalent, or even better, comprehension levels in digital-based reading compared to paper-based reading (see illustration in Fig. 3.1, left panel). To explore this view, we investigated whether the publication date reveals a decreasing advantage of paper in recent years due to greater exposure to technology than in earlier years. If this was the case, with enough experience with digital technologies, readers would be able to overcome any potential detrimental effect on comprehension. In our schematic presentation (Fig. 3.1), we use paper comprehension as the reference level and illustrate potential changes in digital-based comprehension relative to it. Importantly, because we analyse effect sizes rather than objective measures of performance, we cannot know whether this paper-based reference level changes over time. In particular, one could also argue that because new generations may have less exposure to printed texts, paper comprehension will decrease rather than remaining constant. In any of those two cases, the prediction about the evolution of digital-based reading from this perspective is that reading ability on this medium will improve with further experience. Therefore, the advantage of print over digital-based reading will decrease over the years, regardless of the pattern of change in paper comprehension.

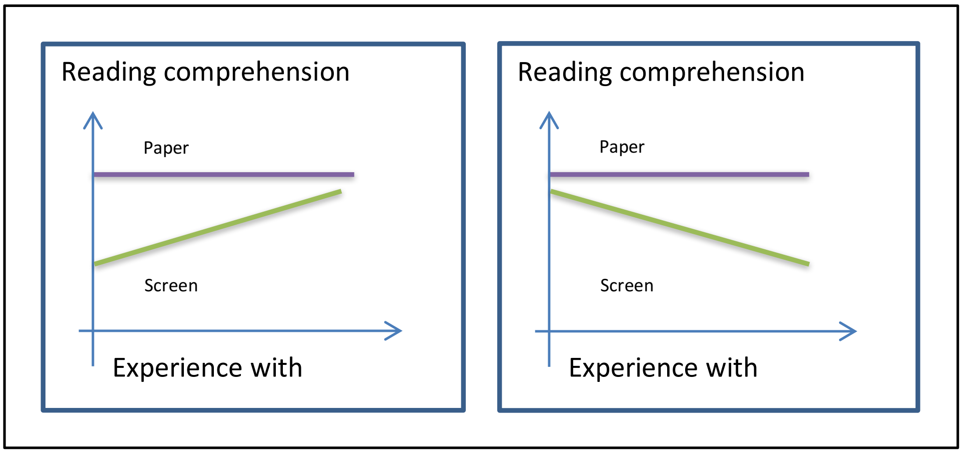

Figure 3.1. Schematic projection of trends for the effect of experience with technology on reading comprehension differences between print and digital devices.

Note. Left panel represents a situation in which more experience with technology reduces the difference between print and digital reading outcomes. Right panel represents a situation in which this potential difference increases over the years.

Several researchers have argued, however, that increasing exposure to technology, with its emphasis on speed and multitasking, may encourage a shallower kind of processing that leads to a decrease in deep comprehension in digital environments (e.g. Lauterman & Ackerman, 2014; Wolf & Barzillai, 2009). Indeed, current evidence supports the claim that mere experience with digital technology does not improve students' comprehension skills, but instead has a detrimental effect (Duncan, McGeown, Griffiths, Stothard, & Dobai, 2015; Pfost, Dörfler, & Artelt, 2013). This view leads to the alternative hypothesis that the paper advantage over digital media increases with time (Fig. 3.1, right panel). If true, this would be a call for researchers, policy-makers, and education professionals to join forces to develop methods to support effective digital-based reading and learning.

3.1.4. Objectives

The aim of this meta-analysis was to gain a broad perspective of empirical studies comparing digital and print reading outcomes. Specifically, we had two objectives:

1) Examine whether the reading medium affects reading comprehension outcomes.

2) Identify moderating factors of the effects of the medium on reading comprehension outcomes.

3.2. Method

3.2.1. Selection criteria of the studies

Studies included in the meta-analysis met the following criteria:

1. The study compares comprehension in paper-based and digital-based reading,

respectively defined as reading texts printed on paper and reading texts displayed on digital screens, including computers, tablets, mobiles phones, and e-readers.

2. Participants read individually and silently.
3. Reading materials are comparable across media in terms of text content, structure, and presence of images. Therefore, specific features of digital environments, such as hyperlinks or web navigation, are not present in the digital-based condition.
4. Participants study in their daily-used language.
5. Participants are a sample from a normative population (i.e., typical development, no reading difficulties, and no cognitive impairments or disorders).
6. The study makes an empirical contribution that includes the results of the comparison (i.e. the paper is not a review or an opinion).
7. The study was published or presented from the year 2000–2017. Formal publication was not required.
8. The report is written in English.
9. The report includes specification of the effect size or sufficient statistical information to calculate it (or this information was provided by the authors following a personal request).
10. The statistical data allow parametric analyses.

3.2.2. Search procedure

Several literature search procedures were used to locate relevant studies and previous reviews. Firstly, some electronic databases were consulted: PsycInfo, Eric, Proquest

Psychology, Web of Science, Scopus (Physical Sciences and Social Sciences & Humanities), dissertation and theses (Proquest), and Google Scholar. The search included the following terms[3]: "("computer reading" OR "online reading" OR "screen reading" OR "digital reading" OR "print reading" OR "paper versus screen" OR "differential test" OR "computer- based testing" OR "computerized testing" OR "computer assisted testing" OR "electronic book" OR "electronic text" OR "media effects" OR "reading medium" OR "mode effect") AND (memory OR comprehension OR retention OR "test performance" OR learning)". These terms were searched as title, abstract, or keywords. As recommended by Card (2012), we complemented the search with additional strategies. Thus, secondly, references included in previous reviews were examined. Thirdly, we approached experts and societies in this area (The Society for Text and Discourse, Society for the Scientific Study of Reading, The European Association for Research on Learning and Instruction, and COST E-READ Action) asking for information about unpublished studies. Fourthly, a forward search was performed using Google Scholar to find studies that cited the works selected. Finally, references from the selected studies were also retrieved. The search ended in May 2017.

The search described above yielded 1,840 records. The selection process from this initial collection is described in Fig. 3.2. We ended up with 54 studies that satisfied all the inclusion criteria. Some studies reported more than one media comparison due to

[3] The study of media effects on reading comprehension has been the focus of several disciplines, including reading research, reading assessment, educational practice, media studies and learning technologies. Each discipline tends to use idiosyncratic words for similar, if not identical, scenarios. For example, the dependent variable in a situation where students read a text and answer comprehension questions is termed "test performance" in the assessment literature, but the term "comprehension scores" is used in the reading literature. Therefore, to avoid leaving out relevant studies from a particular field, we opted to include a broad range of search terms in our query.

considering additional independent factors (e.g., educational level, text genre, digital devices). See the *effect size index* section below for details about the use of these subgroups. The final sample consisted of 76 media comparisons, each contributing an individual effect size. The meta-analysis is based on 171,055 participants. See Appendix A (Table A1 and Table A2) for a detailed distribution of the participants among the studies.

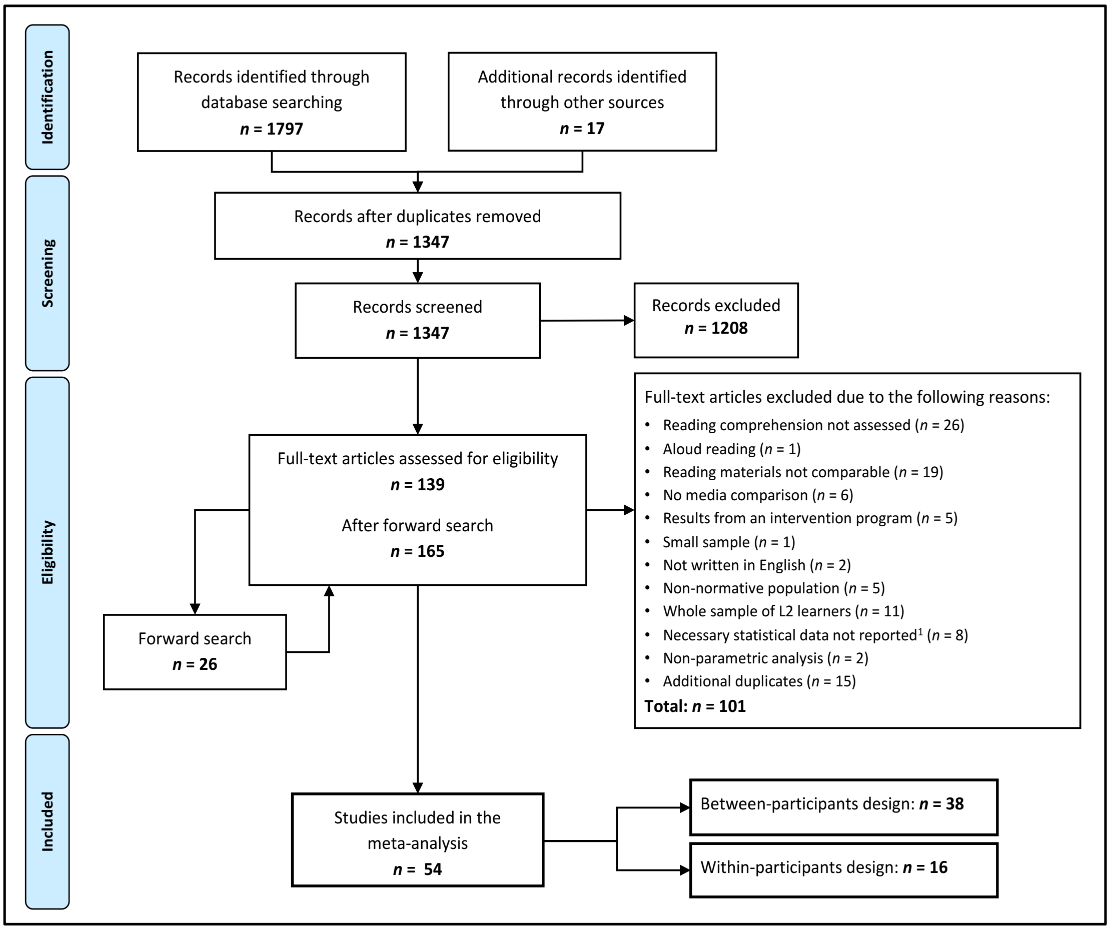

Figure 3.2. Flowchart of the selection process.

Note. [1]Not reported in the study report and not provided by authors following a personal request.

3.2.3. Coding the studies

Several characteristics were coded for each comparison. This allowed for descriptive information and the consideration of moderating variables for the reported effect sizes. When necessary information was not included in the paper for a particular variable, it was coded as "Not reported" (N/r). When available, the following variables were coded:

Substantive variables:

01. *Participants' educational level:* elementary, middle or high school, undergraduates, or graduates and professionals.
02. *Text length:* number of words used in the reading task or other relevant information, such as the number of pages. Once coded, text length was categorized as (a) short (less than 1000 words) or (b) long.
03. *Allowed reading time frame:* (a) free, when reading-time was self-paced by participants, or (b) limited, when time was restricted by experimental instructions.
04. *Type of digital device:* (a) computer (desktop or laptop) or (b) hand-held (tablet, e-reader, or smartphone).
05. *Text genre:* (a) informational, when texts were expository, descriptive or informative, (b) narrative, or (c) mixed, when both genre categories were used in the same task.
06. *Need for scrolling:* whether participants needed to scroll down the texts when reading in digital-based conditions. Coded as (a) yes or (b) no.
07. *Open testing:* whether participants could go back to texts when answering questions. Coded as (a) yes or (b) no.
08. *Type of comprehension:* (a) textual, when reading tasks asked for specific details or shallow level of comprehension; (b) inferential, high-level comprehension, when

tasks required inferences based on parts of the texts, across parts, or involved previous knowledge; or (c) mixed, when tasks required both types of comprehension.

09. *Explicit strategy requirement:* whether participants were prompted or asked to implement a specific strategy in order to promote more in-depth reading, by means of selecting keywords, the use of highlighting or note-taking, or the use of reading strategies promoted by the experimental instructions. Coded as (a) yes or (b) no.

Extrinsic variables:

10. *Publishing status:* (a) published paper, (b) official report, (c) master or PhD thesis, and (d) conference communication.
11. *Year of publication/presentation:* exact year.

Methodological variables:

12. *Sample size:* number of participants.
13. *Sampling method:* (a) probability (some process or procedure that ensures that the different units in the population have equal probabilities of being chosen) or (b) non-probability.
14. *Allocation of participants to media conditions:* (a) random, (b) quasi-random, (c) non-random but matched or controlled, (d) non-random and not controlled, and (e) within-participants design.
15. *Type of reading comprehension test:* (a) standardized/official test or (b) researcher-created task.
16. *Testing medium:* whether participants completed the comprehension test (a) on the same medium used for reading the texts, (b)always on paper, or (c) always on the digital device.

The coding process was conducted by two independent judges, based on a random sample (28%) of the studies included in the meta-analysis. Inter-rater reliability was adequate, showing a Cohen's kappa equal to .89 (minimum=.71, maximum=1) for qualitative variables, and an intra-class correlation (95% CI) yielding absolute agreement for continuous variables (ICC=1). Disagreements were discussed. For transparency and objectivity, a coding manual was developed and is available by request from the last author[4]. A descriptive overview of the studies included is given in the *Results* section and in the Appendix A (Table A1 and Table A2).

3.2.4. The effect size index

The effect size was calculated for each comparison, using means, standard deviations, and sample sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). When the studies used a between-participants design, the standardized mean difference, Hedges' g, was used as the effect size index. This index was defined as the difference between the digital-based (treatment) and paper-based (control) groups' means on the post-test, divided by a pooled within-group standard deviation (Cohen, 1988). In addition, to estimate unbiased effect sizes, the correction factor for small sample sizes proposed by Hedges and Olkin (1985) was used. A positive Hedges' g indicates better comprehension results for the digital-based condition, whereas a negative Hedges' g indicates better outcomes for the paper- based condition.

For studies that used a within-participants design (each participant read on both paper and digital presentations), the standardized mean change index, d_C, was used to

[4] See Appendix A.

estimate the effect sizes. This effect size index is defined as subtracting the mean of the treatment group from the mean of the control group, and then dividing it by the standard deviation of the control group (Botella & Sánchez-Meca, 2015; Morris, 2000). In this case, in order to keep the interpretation of the direction of the mean effect size constant across both datasets (i.e., a positive value indicates better reading outcomes for the digital-based condition and vice versa), we used the digital-based condition as the control group. None of the studies reported the correlation coefficients, and thus, all values were imputed for a conservative estimate (r = .7), as recommended by Rosenthal (1991). As in the previous index, the correction factor for small sample sizes was applied to calculate this effect size index (Hedges & Olkin, 1985).

Finally, as indicated above, some studies reported multiple comparisons. In these cases, the following strategies were applied: a) when the study contained multiple between-participants treatments, the effect size for each subgroup was estimated; b) when there were multiple-treatment groups but they were dependent subgroups, effect sizes and their variances were combined into overall effect sizes and variances for these subgroups; c) if two digital-based groups were compared with the same control group, the sample size for the control group was divided by two to minimize dependence (Higgins & Green, 2011); and d) when the study provided data on multiple outcome measures, effect sizes and variances were averaged to create a single effect size and allow statistical in-dependence of the data (Lipsey & Wilson, 2001). In one case, a combination of strategies *b* and *c* had to be applied due to the existence of three digital-based reading groups.

3.2.5. Statistical analyses

Two separated meta-analyses were performed because it is not recommended to combine studies with between-participants and within-participants designs in one meta-analysis (Lipsey & Wilson, 2001). In each meta-analysis, a weighted mean effect size with its confidence interval (95%) was estimated, and a forest plot was made. Cochran's Q statistic was used to assess the presence of heterogeneity (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006), and I^2 index estimated the proportion of observed variance that is not due to sampling error. Furthermore, the prediction interval was calculated to provide additional context. A random-effects model was used to analyse effect sizes because it is generally regarded as more realistic (Borenstein et al., 2009; Borenstein, Higgins, Hedges, & Rothstein, 2017; Cooper, Hedges, & Valentine, 2009; Huedo-Medina et al., 2006).

Between-study heterogeneity was examined with ANOVAs for qualitative moderators and simple meta-regression for continuous moderators (Borenstein et al., 2009; Cooper et al., 2009), applying the adjustment proposed by Knapp and Hartung (2003). The proportion of variance explained by moderators was estimated by the R^2 index (Raudenbush, 2009).

The normality assumption and outlier detection were assessed by examining the Q–Q normal plot, the Kolmogorov-Smirnov test with the Lilliefors correction, the Shapiro-Wilk test, and the standardized residuals (values greater than 3 in absolute magnitude were considered outliers). When potential outliers were identified, the robust model proposed by Beath (2014) was applied to confirm, removing effect sizes when a probability greater than .9 was found.

Sensitivity analysis was performed to evaluate the robustness of the results. The one-study-removal approach was used to evaluate the impact of each effect size on the mean estimate of the mean effect obtained (Borenstein et al., 2009). Moreover, when calculating the mean effect size for within-participants comparisons, due to the small number of effect sizes, additional methods were used to estimate τ^2 (in particular, the DerSimonian and Laird method with Knapp and Hartung adjustment, the maximum likelihood estimator, and the restricted maximum likelihood estimator). Finally, we also estimated the mean effect sizes, imputing different correlation coefficients (range of values from .10 to .90).

Publication bias was evaluated using Rosenthal's file drawer analysis (Rosenthal, 1979) and Egger's linear regression (Card, 2012), and applying ANOVA to compare the mean effect size of the published versus unpublished studies.

The statistical analyses were conducted using Comprehensive Meta-analysis software Version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2014), R 3.1.1 software with *Metafor* (Viechtbauer, 2010) and *Metaplus* (Beath, 2015) packages, and a Microsoft Excel spreadsheet for computing prediction intervals.

3.3. Results

3.3.1. Descriptive characteristics of the studies

In the final sample (n = 54), 38 studies used a between-participants design. Of these 38 studies, 58 media comparisons (i.e., effect sizes) with 169,524 participants were initially included in the meta-analysis. Note that the majority of these participants (165,778) were from four large-scale studies (Eyre, Berg, Mazengarb, & Lawes, 2017; Lenhard et al., 2017; Pommerich, 2004; Puhan, Boughton, & Kim, 2005; see Appendix A, Table A1). In addition, 16 studies used a within-participants design, providing 18 media comparisons with 1,531 participants. Within our dataset, two studies (Pommerich, 2004; Pomplun, Frey, & Becker, 2002) were included in both the Wang et al. (2007) and Kingston (2008) meta-analyses, mentioned above. Another study (Higgins, Russell, & Hoffman, 2005) was also included in Kingston's work. The remaining studies included in these two meta-analyses did not meet our inclusion criteria.

3.3.1.1. Between-participants studies

Focusing on the substantive variables described in the Appendix A (Table A1), it is worth noting that the majority of the comparisons were conducted with undergraduate students (63.79%), used computers as digital devices (74.13%), included only informational texts (55.17%), and assessed comprehension by means of a mixture of textual and inferential questions (72.41%). In addition, in 44.83% of the comparisons, researchers imposed time constraints for reading the texts. Regarding extrinsic variables, 25 studies (39 effect sizes) were published papers, whereas the remaining 13 studies (17 effect sizes) included PhD dissertations (n = 6), a master thesis (n = 1), conference communications (n = 4), and an official report (n = 1). Moreover, an overview of the between-participants studies shows that 11 studies (16 effect sizes) were published or presented between 2000 and 2010, and 27 studies (42 effect sizes) between 2011 and 2017. Finally, regarding the methodological variables, 98.27% of the comparisons were from studies that recruited the sample through a non-probability sampling method, and 74.14% reported a randomized group allocation of participants. Researcher-created tasks were used in approximately 63.79% of the comparisons (see Appendix A, Table A1, for additional information).

Finally, it is worth noting that several studies did not report information about some of the coded variables. However, they were included in the dataset whenever the information provided allowed us to calculate effect sizes because our purpose was to include a sample of studies in the meta-analysis that was as representative as possible.

3.3.1.2. Within-participants studies

The within-participants studies included are described in the Appendix A (Table A2). Regarding substantive variables, a majority of the 18 comparisons reported that they were conducted with undergraduates (55.55%), used computers for digital-based reading (55.55%), used informational texts (61.11%), and assessed comprehension by means of a mixture of textual and inferential questions (55.55%). In relation to reading time, five comparisons imposed time constraints. Focusing on extrinsic variables, this dataset consisted of 11 published studies (13 effect sizes), a PhD dissertation, a bachelor thesis, and three conference communications (in all, 5 effect sizes from unpublished studies). Only four studies were reported before 2011. With regard to methodological variables, all the studies recruited the sample through a non-probability method, and eleven comparisons were conducted using researcher-created tasks.

3.3.2. The mean effect size, heterogeneity, and sensitivity analyses

Before calculating the mean effect size, preliminary analyses were conducted to identify outliers and verify normality of the sample. Two effect sizes were identified as possible outliers (Duran, 2013; Nishizaki, 2015; see Appendix A, Table A1) by examining standardized residuals (values > 3), the Q–Q normal plot, and the Kolmogorov-Smirnov test with the Lilliefors correction (p = .02) in the between-participants dataset. The robust model was applied to further analyse these potential outliers, with both obtaining probabilities greater than 0.90. Therefore, they were removed from posterior analyses, and so the final sample of between-participants studies included 56 effect sizes. After removing outliers, effect sizes were normally distributed (p = .40).

When examining the within-participants dataset, no effect size was identified as an outlier, and so the initial 18 effect sizes were all included in the analysis. The Shapiro-Wilk normality test (p = .52) indicated that the dataset was normally distributed.

3.3.2.1. Media effect in between-participants designs

As explained above, comprehension in paper-based reading groups was used as the baseline. Therefore, negative values indicate that reading outcomes from digital-based devices were lower than their respective paper-based groups. The mean effect size of the sample was significant (Hedges' $g = -0.21$; 95% CI: −0.28, −0.14; k = 56), revealing an advantage of paper-based reading over digital-based reading. An overview of the effect sizes can be seen in Fig. 3.3, which provides a graphical representation of the estimated results of each reading media comparison. Each result is represented by a blue line with a dot in the centre. The dot indicates the value of the effect size (note the vertical lines marking values from −2 to 2), and the line that emerges from both sides of the dot represents the confidence interval. The longer the line, the larger the confidence interval. Lines that do not reach the zero value indicate significant effect sizes.

Regarding the variability of the effect sizes, the heterogeneity between individual effect sizes was medium-high (I^2 = 72.24) and statistically significant (Q = 208.96, p < .001). The prediction interval was −0.56 to 0.14, and so it was expected that the true effect size would fall in this range in 95% of all populations. Hence, the effects are large in some populations, but moderated and trivial in other populations. The wide range of effects calls for further analyses to examine potential moderating factors that would shed light on sources of differences among the studies. Thus, analyses were conducted to examine effects of substantive, extrinsic, and methodological variables. The results are reported below.

	Hedges' *g*	Lower	Upper	*p*
Ackerman & Lauterman, 2012 (Exp. 1)a	-1,17	-1,82	-0,52	0,00
Ben-Yehudah & Eshet-Alkalai, 2014a	-0,80	-1,39	-0,21	0,01
Ackerman & Goldsmith, 2011 (Exp. 2)	-0,77	-1,23	-0,30	0,00
Nishizaki, 2015 (Exp. 1)b	-0,67	-1,29	-0,04	0,04
Beach, 2008b	-0,66	-1,26	-0,05	0,03
Nishizaki, 2015 (Exp 2)	-0,64	-1,08	-0,19	0,00
Chen et al., 2014a	-0,64	-1,26	-0,01	0,05
Ackerman & Lauterman, 2012 (Exp. 2)	-0,63	-1,09	-0,18	0,01
Jones et al., 2005	-0,58	-1,15	-0,00	0,05
Wästlund et al., 2005	-0,56	-1,03	-0,09	0,02
Norman & Furnes, 2016 (Exp. 2)	-0,52	-1,08	0,03	0,06
Mayes et al., 2001	-0,50	-1,12	0,12	0,11
Hongler, 2015b	-0,47	-1,15	0,22	0,18
Chen et al., 2014b	-0,44	-1,05	0,18	0,16
Mangen et al., 2013	-0,44	-0,92	0,05	0,08
Eyre et al., 2017a	-0,38	-0,44	-0,33	0,00
Eyre et al., 2017b	-0,38	-0,44	-0,32	0,00
Bartell et al., 2006	-0,37	-0,62	-0,11	0,00
Taylor, 2011b	-0,36	-1,01	0,30	0,28
Simian et al., 2016	-0,35	-0,76	0,07	0,10
Grimshaw et al., 2007a	-0,33	-0,87	0,22	0,24
Lenhard et al., 2017	-0,32	-0,40	-0,25	0,00
Green et al., 2010	-0,31	-0,84	0,22	0,25
Ben-Yehudah & Eshet-Alkalai, 2014b	-0,28	-0,84	0,29	0,33
Higgins et al., 2005b	-0,25	-0,64	0,14	0,21
Margolin et al., 2013b	-0,23	-0,73	0,28	0,38
Seehafer, 2014	-0,21	-0,69	0,26	0,38
Norman & Furnes, 2016 (Exp. 1)b	-0,21	-0,87	0,45	0,54
Lauterman & Ackerman, 2014 (Exp. 2)	-0,20	-0,64	0,25	0,39
Hongler, 2015a	-0,19	-0,87	0,49	0,59
Lauterman & Ackerman, 2014 (Exp. 1)	-0,16	-0,58	0,27	0,47
Pommerich, 2004 (Exp. 1)	-0,14	-0,23	-0,05	0,00
Norman & Furnes, 2016 (Exp. 1)a	-0,13	-0,80	0,54	0,70
Grimshaw et al., 2007b	-0,13	-0,65	0,39	0,63
Kaufman & Flanagan, 2016	-0,10	-0,55	0,34	0,64
Johnson, 2013	-0,06	-0,32	0,19	0,63
Higgins et al., 2005a	-0,05	-0,44	0,34	0,80
Daniel & Woody, 2013	-0,05	-0,55	0,46	0,86
Pommerich, 2004 (Exp. 2)a	-0,04	-0,14	0,07	0,50
Pommerich, 2004 (Exp. 2)b	-0,03	-0,13	0,08	0,58
Burkley, 2013	-0,02	-0,69	0,65	0,95
Beach, 2008a	-0,01	-0,77	0,75	0,98
Chen, 2015	0,00	-0,41	0,41	1,00
Niccoli, 2015	0,01	-0,24	0,27	0,91
Wells, 2012	0,02	-0,31	0,35	0,92
Ackerman & Goldsmith, 2011 (Exp. 1)	0,03	-0,44	0,49	0,91
Puhan et al., 2005	0,03	-0,06	0,11	0,51
Margolin et al., 2013a	0,04	-0,46	0,54	0,87
Porion et al., 2016	0,05	-0,41	0,51	0,83
Ackerman & Lauterman, 2012 (Exp. 1)b	0,06	-0,56	0,67	0,86
Morineau et al., 2005	0,08	-0,53	0,69	0,80
Connell et al., 2012b	0,11	-0,32	0,54	0,61
Connell et al., 2012a	0,11	-0,31	0,54	0,60
Taylor, 2011a	0,26	-0,40	0,92	0,43
Aydemir et al., 2013	0,39	-0,11	0,89	0,13
McCrea-Andrews, 2014	0,61	-0,04	1,27	0,07
Mean effect	-0,21	-0,28	-0,14	0,00

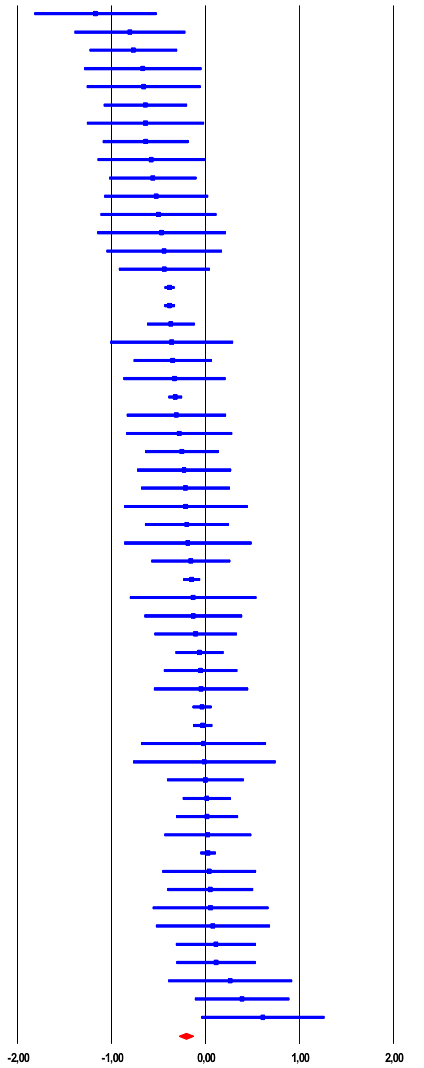

Figure 3.3. Forest plot of reading media effect sizes on reading comprehension from studies using between-participants designs.

Note. Letters after the publication year differentiate several comparisons from the same study. Note that comparisons reported in the studies could have been recoded in the meta-analysis (see *Method* section). Please note that negative values indicate better outcomes for paper-based reading.

3.3.2.2. Sensitivity analyses for between-participants comparisons

The one-study-removal method (Borenstein et al., 2009) showed that effect sizes fell between Hedges' $g = -0.22$ and -0.20 ($p < .001$) and did not substantially affect the mean effect size, indicating a significant advantage of paper-based reading in all cases. Special attention should be paid to the four large-scale studies mentioned above (Eyre et al., 2017; Lenhard et al., 2017; Pommerich, 2004; Puhan et al., 2005). Given that their large samples yielded a small confidence interval for their effect sizes, their influence on the overall effect could skew the results. However, excluding these studies altogether (7 effect sizes), the mean effect size was Hedges' $g = -0.22$ ($p < .001$), which means they did not bias the overall effect of the reading media. Finally, given that we included "grey literature" (unpublished studies) in our meta-analysis, we repeated the meta-analysis without these studies in order to make sure that their inclusion does not compromise research quality. The mean effect size was Hedges' $g = -0.19$ (95% CI: −0.27, −0.11; $k = 38$) when excluding all the unpublished studies (i.e., official reports, conference communications, and dissertations) and Hedges' $g = -0.20$ (95% CI: −0.28, −0.13; $k = 51$) when only excluding the conference communications. Thus, "grey literature" did not substantially affect the overall mean effect size in this dataset.

3.3.2.3. Media effect in within-participants designs

The mean effect size of this sample of studies was also significant, and it replicated the advantage of paper-based reading over digital-based reading ($d_C = -0.21$; 95%; CI: −0.37, −0.06; $k = 18$). Fig. 3.4, similarly to Fig. 3.3, presents an overview of the effect sizes included in the dataset of studies that used a within-participants design.

As in between-participants studies, heterogeneity of the effect sizes was high (I^2

= 89.88; Q = 167.94, p < .001), with the prediction interval ranging from −0.90 to 0.47. Nevertheless, analyses of moderators were not performed in this dataset, due to the small number of effect sizes, and this should be taken into account when interpreting the results.

	d_c	Lower	Upper	p
Kim & Huynh, 2008	-0,75	-0,84	-0,67	0,00
Kim & Kim, 2013	-0,63	-0,80	-0,46	0,00
Jeong, 2012	-0,40	-0,62	-0,18	0,00
Kretzschmar et al., 2013a	-0,37	-0,64	-0,09	0,01
Heij & van der Meij, 2014	-0,36	-0,77	0,05	0,09
Sackstein et al., 2015b	-0,32	-0,75	0,12	0,16
Singer & Alexander, 2017	-0,21	-0,37	-0,05	0,01
Baker, 2010	-0,21	-0,36	-0,05	0,01
Kretzschmar et al., 2013b	-0,21	-0,55	0,13	0,23
Delgado & Salmerón, 2017	-0,20	-0,38	-0,01	0,04
Rasmusson, 2015	-0,16	-0,30	-0,01	0,03
Thompkins et al., 2016	-0,13	-0,33	0,07	0,20
Bansi et al., 2016	-0,12	-0,40	0,17	0,42
Sackstein et al., 2015a	-0,10	-0,31	0,10	0,33
Hermena et al., 2017	0,00	-0,31	0,31	1,00
Kerr & Symons, 2006	0,08	-0,12	0,28	0,45
Pomplun et al., 2002	0,09	-0,02	0,19	0,10
Liang & Huang, 2013	0,19	-0,14	0,51	0,26
Mean effect	-0,21	-0,37	-0,06	0,01

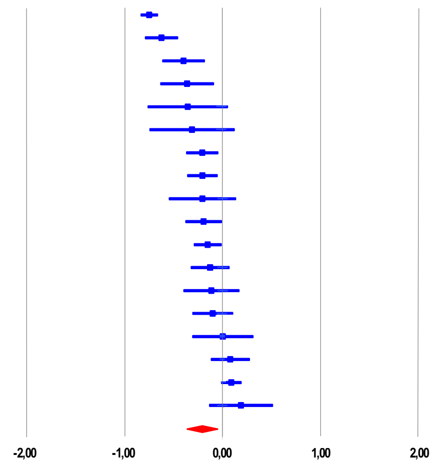

Figure 3.4. Forest plot of reading media effect sizes on reading comprehension from studies using within-participants designs.

Note. Letters after the publication year differentiate several comparisons from the same study. Note that comparisons reported in the studies could have been recoded in the meta-analysis (see *Method* section). Please note that negative values indicate better outcomes for paper-based reading.

3.3.2.4. Sensitivity analyses for within-participants comparisons

One-study-removal analysis (Borenstein et al., 2009) indicated that effect sizes fell between d_C = −0.18 and −0.24 (p < .001), and were again significant, showing an advantage of paper-based reading in all cases. Additional results from Knapp and Hartung's adjustment of the DerSimonian and Laird estimator (d_C = −0.21; 95% CI: −0.33, −0.09; k = 18), the maximum likelihood approach (d_C = −0.22; 95% CI: −0.33, −0.10; k = 18), and the restricted maximum likelihood method (d_C = −0.22; 95% CI: −0.34, −0.10; k = 18) were also consistent. Moreover, a sensitivity analysis imputing different correlation coefficients (range of values from 0.10 to 0.90) was carried out. The findings were essentially identical (the largest difference between mean effect sizes was smaller than 3%) and revealed that the meta-analysis result was robust. Consequently, the result reported was based on a correlation of .70, as recommended by Rosenthal (1991). In addition, we also examined whether the inclusion of unpublished studies affected the overall effect of the reading media in this dataset. Thus, the mean effect size was d_C = −0.22 (95%; CI: −0.42, −0.13; k = 13) when excluding all the unpublished studies, and d_C = −0.23 (95%; CI: −0.41, −0.04; k = 15) when only excluding the conference communications. Therefore, "grey literature" did not affect the overall mean effect size in within-participants studies either.

3.3.3. Publication bias

3.3.3.1. Publication bias for between-participants comparisons

The risk of publication bias was examined with three different methods. First, results from Classic fail safe-N analysis indicated that 1,727 null effect sizes would be necessary to nullify the mean effect size of the medium. This value meets Rosenthal's criterion (5k + 10), which sets 290 as the minimum for this dataset. Second, Egger's linear regression indicated a non-significant publication bias (p = .39). Finally, an ANOVA revealed that the mean effect sizes from published versus unpublished studies were not statistically different (Q_B (1, 54) = 0.14, p = .71). All these results suggested that there was no publication bias.

3.3.3.2. Publication bias for within-participants comparisons

In this dataset, Classic fail Safe-N analysis indicated that 475 null effect sizes would be necessary to nullify the mean effect size of the media, which again was a higher value than Rosenthal's criterion ($5k + 10 = 100$). Additionally, Egger's linear regression yielded a non-significant publication bias (p = .20), and an ANOVA between published and unpublished studies showed no significant differences (Q_B (1, 16) = 0.02, p = .90. Likewise, these three indicators suggested no risk of publication bias.

3.3.4. Moderating variables in between-participants comparisons

In the following analyses, we considered potential moderating variables, grouped by substantive, extrinsic, and methodological variables, for media effects on reading outcomes among the between-participants studies. As mentioned above, some studies lacked the necessary information about some of these variables, and so they were not included in the respective moderator analyses.

Substantive variables

We conducted an ANOVA for each substantive variable considered. These analyses indicated significant moderating effects of the allowed reading time frame (i.e., limited by task constraints vs. self-paced by participants) and text genre (i.e., informational texts vs. narrative texts vs. a combination of both genres). No moderating effects were found for educational level, text length, type of digital device, need for scrolling, open testing, or type of comprehension because Q_B values were not significant in all these cases (see Table 3.1). Examination of the reading time frame showed that comparisons in studies with time constraints yielded a significantly larger (Q_B (1, 45) = 4.12, p = .04) print advantage (Hedges' $g = -0.26$) than comparisons in studies in which participants were allowed to self-pace their reading (Hedges' $g = -0.09$). Thus, although there is an overall advantage of print over digital devices, the difference is larger with time constraints than with self-paced reading, which explains 5% of the mean effect size variance.

The moderator factor of text genre revealed a significant effect, explaining 31% of the mean effect size variance. Comparisons conducted with informational texts or a combination of informational and narrative texts showed significant mean effect sizes favouring paper-based reading over digital-based reading (Hedge's $g = -0.27$ and -0.30, respectively), whereas comparisons conducted only with narrative texts showed no effect of media (Hedge's $g = 0.01$) (see Table 3.1).

Two variables are worth mentioning, even though their moderating effects did not reach significance. The advantage of paper- based reading was significant when studies used computers (Hedges' $g = -0.23$, $p < .001$), but not when they used hand-held devices (Hedges' $g = -0.12$, $p = .11$). Similarly, the need for scrolling as a feature of digital-based reading resulted in a significant advantage of paper-based reading (Hedges' $g = -0.25$, $p < .001$), whereas the media effect was marginal and numerically smaller when scrolling was not necessary (Hedges' $g = -0.13$, $p = .06$) (see Table 3.1).

Finally, due to the small number of comparisons where in-depth reading was prompted by means of an explicit strategic requirement ($k = 5$), the moderating effect of this variable was not examined.

Extrinsic variables

As reported above, the ANOVA with publishing status was not a significant moderator, as indicated by the Q_B value (see Table 3.2). However, a meta-regression analysis revealed that the date of publication or presentation of the studies has a significant moderating effect on the mean effect size of the media. The advantage of paper-based reading over digital-based reading increased since 2000, as hypothesised in the right panel of Fig. 3.1. The beta coefficient of -0.01 ($Q_R = 4.95$, $p = .03$) indicates that the effect size favouring paper-based reading increased by 0.01 points a year, explaining 64% of the mean effect size variance (see Table 3.3).

Table 3.1. *One-way analysis of variance of substantive variables on mean effect sizes for reading media from the studies using between-participants designs.*

			Mean effect sizes				
Variable[1]	**Categories**	***k***	**Hedges’ *g***	**95% CI**	**Q_B(df)**	**Q_w(df)**	**R^2**
Participants’ educational level[2]					2.33(2)	131.33(49)***	.00
	Grades 1 to 6	8	-0.19	[-0.35, -0.03]			
	Grades 7 to 12	8	-0.15	[-0.29, -0.02]			
	Undergraduates	36	-0.28	[-0.38, -0.18]			
Text length					0.14(1)	142.36(47)***	.00
	Short	22	-0.25	[-0.34, -0.16]			
	Long	26	-0.22	[-0.33, -0.11]			
Reading time-frame					4.12(1)*	185.17(45)***	.05
	Self-paced	20	-0.09	[-0.22, 0.05]			
	Limited	27	-0.26	[-0.35, -0.16]			
Digital device					1.55(1)	194.95(54)***	.02
	Computer	42	-0.23	[-0.31, -0.15]			
	Hand-held	14	-0.12	[-0.27, 0.03]			
Text genre					7.00(2)*	74.21(48)**	.31
	Informational	34	-0.27	[-0.36, -0.18]			
	Narrative	7	0.01	[-0.20, 0.20]			
	Mixed	10	-0.30	[-0.40, -0.21]			
Need for scrolling					1.99(1)	133.40(47)***	.00
	No	12	-0.13	[-0.27, 0.01]			
	Yes	37	-0.25	[-0.33, -0.16]			
Open testing					1.21(1)	183.46(47)***	.00
	No	33	-0.26	[-0.37, -0.16]			
	Yes	16	-0.18	[-0.29, -0.07]			
Type of comprehension[3]					0.14(1)	153.99(51)	.00
	Textual	9	-0.26	[-0.47, -0.04]			
	Mixed + Inferential	44	-0.21	[-0.29, -0.14]			

Note. k: number of effect sizes. Hedges’ ***g*****:** mean effect size. Q_B: between-categories Q statistic. Q_W: within-categories Q statistic. R^2: Proportion of total between-comparison variance explained. [1]Non-reported values for each variable were not included in these analyses. [2]Due to the small number of effect sizes, the category “Graduates or professionals” (k = 3) was not included in this analysis. [3]Due to the small number of effect sizes, comparisons that examined only inferential comprehension (k = 3) were included in the same group as those that examined both types of comprehension. *$p < .05$. **$p < .01$. ***$p < .001$.

Methodological variables

Four methodological variables were tested to examine their possible influence on the media effect. They were sample size, method of allocating participants to media conditions, the type of reading comprehension test, and the testing medium. Results revealed that none of these four methodological variables had a significant moderating effect, as indicated by the Q_B and Q_R values (see Tables 3.2 and 3.3). The sampling method variable was not analysed due to lack of variability (See Appendix A, Table A1).

Table 3.2. *One-way analysis of variance of moderating effect of extrinsic and methodological variables on mean effect sizes for reading media from the studies using between-participants designs.*

			Mean effect sizes				
Variable[1]	**Categories**	***k***	**Hedges' *g***	**95% CI**	**Q_B(df)**	**Q_w(df)**	**R^2**
Publishing status					0.14(1)	186.47(54)***	.00
	Published	39	-0.22	[-0.31, -0.13]			
	Unpublished	17	-0.19	[-0.31, -0.07]			
Group allocation[2]					0.90(2)	167.33(49)***	.00
	Random	44	-0.20	[-0.28, -0.12]			
	Non-random	7	-0.28	[-0.46, -0.12]			
Type of reading comprehension test					0.01(1)	200.15(54)***	.00
	Standard/official	22	-0.21	[-0.31, -0.11]			
	Researcher-created	34	-0.21	[-0.32, -0.11]			
Testing medium					1.11	180.06(45)***	.00
	Same for reading	27	-0.26	[-0.35, -0.17]			
	Always on paper	20	-0.17	[-0.31, -0.03]			

Note. k: number of effect sizes. Hedges' *g:* mean effect size. Q_B: between-categories Q statistic. Q_W: within-categories Q statistic. R^2: Proportion of total between-comparison variance explained. [1]The variable sampling method was not included in the analyses due to lack of variability. [2]Due to the small number of effect sizes categories "Non-random but controlled" (k = 3) and "Non-random not controlled" (k = 4) were combined ("Non-random"). ***p < .001

Table 3.3. *Meta-regression analysis of moderating effect of sample size and date of publication on mean effect sizes for reading media from the studies using between-participants designs.*

Variable	k	b	Q_R	Q_E	R^2
Sample size	56	-0.00	3.11	201.59***	.42
Date of publication	56	-0.01	4.95*	201.59***	.64

Note. k: number of effect sizes. *b:* unstandardized regression coefficient. Q_R: statistical test of between-comparison effects. Q_E: statistical test of between-comparison homogeneity of the effect sizes. R^2: Proportion of total between-comparison variance explained. *$p < .05$. ***$p < .001$.

3.4. Discussion

This study sought to address an issue of great importance in education and work-related contexts, namely, whether and under what conditions media have an effect on reading comprehension. The strong appeal of digital-based assessment and learning environments has led many educational systems to adopt them. As findings from the current work reveal, however, digital environments may not always be best suited to fostering deep comprehension and learning. The straightforward conclusion is that providing students with printed texts despite the appeal of computerized study environments might be an effective direction for improving comprehension outcomes. However, given the unavoidable inclusion of digital devices in our contemporary educational systems, more work must be done to train pupils on dealing with performing reading tasks in digital media, as well as to understand how to develop effective digital learning environments.

The results of the two meta-analyses in the present study yield a clear picture of screen inferiority, with lower reading comprehension outcomes for digital texts compared to printed texts, which corroborates and extends previous research (Kong et al., 2018; Singer & Alexander, 2017b; Wang et al., 2007). These results were consistent across methodologies and theoretical frameworks.

Although the effect sizes found for media (−0.21) are small according to Cohen's guidelines (1988), it is important to interpret this effect size in the context of reading comprehension studies. During elementary school, it is estimated that yearly growth in reading comprehension is 0.32 (ranging from 0.55 in grade 1, to 0.08 in grade 6) (Luyten, Merrel, & Tymms, 2017). Intervention studies on reading comprehension yield a mean effect of 0.45 (Scammacca, Roberts, Vaughn, & Stuebing, 2015). Thus, the effects of media are relevant in the educational context because they represent approximately 2/3 of the yearly growth in comprehension in elementary school, and 1/2 of the effect of remedial interventions.

Our investigation of moderating factors indicated that the advantage of paper-based reading is significantly larger when a reading time limit is imposed, compared to self-paced reading. Such advantage is consistent across studies using informational texts (or a mix of informational and narrative), but no media effect is found when the studies used only narrative texts. In addition, the advantage of print reading significantly increased from 2000 to 2017. Furthermore, although they did not reach significance, the results suggest stronger media differences on computers than on hand-held devices, as well as disadvantages of digital texts that require scrolling. Finally, the results indicate that media differences do not vary according to the remaining substantive factors: age group (educational level), text length, type of comprehension assessed, or the option to revise the text to answer the questions; extrinsic factors: sample size and publishing status; or methodological factors: type of test, group allocation, and testing medium.

We discuss below the implications of the findings. In particular, how the screen inferiority effect is related to the reading practices of new generations, to theories of self-regulated learning, and to the genre of the reading materials. We then identify some of the limitations of the study and conclude by discussing several educational implications of our results.

3.4.1. Media effect and new generations

The adoption of new media practices often involves activating a set of cognitive processes appropriate for taking full advantage of the media. For children growing up surrounded by digital technologies, skills such as the ability to search and navigate, read critically, and multitask are essential (e.g. Salmerón, García, & Vidal-Abarca, 2018). Such skills place demands on attention and executive processes that may not be fully developed in children and adults reading digital texts. If simply being exposed to digital technologies were enough to gain these skills, then we would expect an increasing advantage of digital reading, or at least decreasing screen inferiority over the years. Contrary to this assumption, however, our results indicate that the screen inferiority effect has increased in the past 18 years, and that there were no differences in media effects between age groups. These surprising findings suggest that we cannot idly wait for screen inferiority to disappear as children are exposed to digital devices earlier and earlier in their lives, as adults gain more experience with the technology, or as technology improves. The data suggest that screen inferiority is a major challenge across age groups that becomes more severe as the presence of technology increases.

3.4.2. Media effect and time frames for learning

Our results do not address the cause of this persistent screen inferiority, but they provide evidence that people adopt a shallower processing style in digital environments (e.g. Lauterman & Ackerman, 2014; Wolf & Barzillai, 2009). The increase in media differences as technology becomes more integrated into our lives may be related to poorer quality of attention (Courage, 2017), where deep immersion in the text is challenged (e.g. Mangen & Kuiken, 2014). The Shallowing Hypothesis suggests that because the use of most digital media consists of quick interactions driven by immediate rewards (e.g. number of "likes" of a post), readers using digital devices may find it difficult to engage in challenging tasks, such as reading comprehension, requiring sustained attention (Annisette & Lafreniere, 2017). According to this perspective, the more people use digital media for these shallow interactions, the less they will be able to use them for challenging tasks. Such arguments are consistent with negative correlations reported between the frequency of digital media use and text comprehension in adolescents (Duncan et al., 2015; Pfost et al., 2013), and they suggest that we should be cautious about the introduction of digital reading in classrooms.

A relevant moderator found for the screen inferiority effect was time frame. This finding sheds new light on the mixed results in the existing literature. Consistent with the findings by Ackerman and Lauterman (2012) with lengthy texts, mentioned above, Sidi et al. (2017) found that even when performing tasks involving reading only brief texts and no scrolling (solving challenging logic problems presented in an average of 77 words), digital-based environments harm performance under time pressure conditions, but not under a loose time frame. In addition, they found a similar screen inferiority when solving problems under time pressure and under free time allocation, but framing the task as preliminary rather than central. Thus, the harmful effect of limited time on digital-based work is not limited to reading lengthy texts. Moreover, consistently across studies, Ackerman et al. found that people suffer from greater overconfidence in digital-based reading than in paper-based reading under these conditions that warrant shallow processing. Sidi et al. (2017) explained that time pressure and framing the task as preliminary both justify shallow processing, which has a stronger effect in digital environments where people are used to quick and shallow tasks (e.g., Facebook, chats; see also Lauterman & Ackerman, 2014). These empirical findings support Annisette and Lafreniere's (2017) Shallowing Hypothesis, which had previously been based on self-reports.

Our findings call to extend existing theories about self-regulated learning (see Boekaerts, 2017, for a review). Effects of time frames on self-regulated learning have been discussed from various theoretical approaches. First, a metacognitive explanation suggests that time pressure encourages compromise in reaching learning objectives (Thiede & Dunlosky, 1999). Second, time pressure has been associated with cognitive load. Some studies found that time pressure increased cognitive load and harmed performance (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). However, others suggested that it can generate a germane ("good") cognitive load by increasing task engagement (Gerjets & Scheiter, 2003). In these theoretical discussions, the potential effect of the medium in which the study is conducted has been overlooked. We see the robust finding in the present meta-analyses about the interaction between the time frame and the medium as a call to theorists to integrate the processing style adapted by learners in specific study environments into their theories.

The finding in this meta-analysis that most media effects come from tasks performed under limited time frames should be taken into account by designers of admission exams and educators. The disadvantage of digital-based reading would be especially critical if not all the examinees are tested in the same medium. Moreover, this could also be an influential factor even when they are all examined by means of digital tests, because of individual differences in adapting to the digital media. For instance, Lauterman and Ackerman (2014) found differences in media effects on learning outcomes based on people's media preference. Clearly, additional individual difference should be considered. Thus, digital exams outcomes probably reflect not only the knowledge or skill at hand, but also such digital-specific competencies.

An encouraging finding from Lauterman and Ackerman (2014) and Sidi et al. (2017) is that simple methodologies (e.g., writing keywords summarizing the text, framing the task as central) that engage people in in-depth processing make it possible to eliminate screen inferiority, in terms of both performance and overconfidence, even under a limited time frame. Together, these findings strongly suggest that pedagogy should play a significant role in identifying individual differences and guiding students to develop skills they miss that support a thoughtful approach to digital information, even when the task design seems to indicate the legitimacy of shallow processing.

3.4.3. Media effect and text genre

The text genre was another variable that moderated media effects. On the one hand, the paper-based reading advantage was consistent across studies using informational texts, or a mix of informational and narrative texts. On the other hand, studies using only narrative texts showed no effect of media on comprehension. Comprehending informational texts, compared to narratives, requires higher level processing, such as using complex academic vocabulary and structures, and these texts are less connected to real world knowledge, which makes them harder to comprehend (Graesser & McNamara, 2011). Thus, our finding may also point to the Shallowing Hypothesis as an explanation. Nevertheless, this result must be interpreted with caution due to the small number of comparisons that used only narrative texts. In addition, among the included studies that directly compared text genre and reading medium, only Simian et al. (2016) reported a significant interaction between these variables, revealing a positive effect of print-based reading only on informational texts, whereas two studies found no effect of text genre (Margolin, Driscoll, Toland, & Kegler, 2013; Rasmusson, 2015).

3.4.4. Additional potential moderators of media effects

Future research should aim to identify other variables that may interact with media effects. In particular, moderators with effects that approached significance deserve further consideration (see Table 3.1), such as the influence of the type of device. It is important to determine whether screen inferiority is limited to desktop computers and eliminated when using hand-held devices. If this proves to be the case, it would be important to understand what cognitive processes could allow media equivalence on hand-held devices. Of the three studies included in this meta-analysis that specifically examined differences among digital devices (Chen, Cheng, Chang, Zheng, & Huang, 2014; Hongler, 2015; Margolin et al., 2013), only Chen et al. (2014) found an interaction with media, reporting a negative impact of digital reading only on computers.

In addition, the need for scrolling was found to be a possible obstacle to comprehension during digital reading. Among the studies included in the meta-analysis, Pommerich (2004) and Higgins et al. (2005) found that participants who read non-scrolling digital texts outperformed those who read scrolling texts, although the differences were not significant. These studies, however, were performed more than a decade ago. Nonetheless, scrolling may add a cognitive load to the reading task by making spatial orientation to the text more difficult for readers than learning from printed text. One of the questions about the scrolling findings is whether the effect of scrolling is related to longer texts or some other artefact of mouse use while reading, although text length was not found to be a moderating factor in our meta-analyses.

3.4.5. Limitations

We would like to call attention to some limitations in our meta-analyses. First, ten studies that met the inclusion criteria could not be included due to lack of necessary statistical data ($n = 8$) or non-normal distributions ($n = 2$).

Moreover, the effect sizes included in the meta-analyses showed high heterogeneity. The moderators considered captured some of this variance, but there is clearly unexplained variance. Consequently, additional factors potentially influencing the results could be affecting the mean effect size. In particular, factors related to research methods (e.g., the reliability of the testing tools) or to sample characteristics (e.g., SES or degree of use of digital texts for learning purposes) could be considered. These factors were missing from most of the reports we included in our meta-analyses. Therefore, we encourage researchers to investigate these possible moderators and describe their methods and samples in detail in future publications.

In addition, the interpretation of how the effect of reading media changes over generations was based on the studies' publication dates. Clearly, using the date as indicator of generation is simplistic and may affect several aspects (e.g., research methods may change throughout the years). In particular, we considered it relevant to examine how different age groups interact with the publication date. However, the distribution of age groups over the years was not broad enough to allow reliable analysis of this possible effect in our dataset. Thus, we recommend considering how different factors interact with the year of publication.

Finally, given that our purpose was to isolate the effect of media, per se, on reading outcomes, we excluded digital affordances (except for scrolling) such as hypertext reading or navigation through webpages. Their effect on reading comprehension is still an open question that warrants further research efforts.

3.5. Conclusions

In conclusion, it is clear that digital-based reading is an unavoidable part of our daily lives and an integral part of the educational realm. Although the current results suggest that paper-based reading should be favoured over digital-based reading, it is unrealistic to recommend avoiding digital devices. Nevertheless, ignoring the evidence of a robust screen inferiority effect may mislead political and educational decisions, and even worse, it could prevent readers from fully benefiting from their reading comprehension abilities and keep children from developing these skills in the first place. Thus, we call on researchers to consider how to guide students and exam takers in dealing with digital tasks such as admission tests (e.g., SAT and GMAT), tasks in work contexts, and school-related tasks that are very often performed with informational texts and under limited time frames. In particular, an important conclusion from our analysis is that there are predictable conditions that seem to allow media equivalence. It is important to appreciate these conditions, examine their validity for the task at hand, and use them whenever possible and relevant. We hope our meta-analysis will guide evidence-based decisions by policy makers and point designers and researchers toward conditions that support effective digital-based reading.

References

* The complete list of the references included in the meta-analysis can be found in Appendix A.

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied, 17*, 18-32. https://doi.org/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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

American Speech-Language-Hearing Association (2015). *Parent pooling: Better hearing and speech month.* http://www.asha.org/uploadedFiles/BHSM-Parent-Poll.pdf.

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154-158. https://doi.org/10.1016/j.paid.2016.02.043

Baron, N. S., Calixte, R. M., & Havewala, M. (2017). The persistence of print among university students: An exploratory study. *Telematics and Informatics, 34*, 590-604. https://doi.org/10.1016/j.tele.2016.11.008

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 570-585. https://doi.org/10.1037/0278-7393.33.3.570

Beath, K. J. (2014). A finite mixture method for outlier detection and robustness in meta-analysis. *Research Synthesis Methods, 5*, 285-293. https://doi.org/10.1002/jrsm.1114

Beath, K. J. (2015). *Metaplus: Robust meta-analysis and meta-regression, 2016.* R package version 0.7-7. Boekaerts, M. (2017). Cognitive load and self-regulation: Attempts to build a bridge. *Learning and Instruction, 51*, 90-97. https://doi.org/10.1016/j.learninstruc.2017.07.001

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis.* Chichester, UK: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2014). *Comprehensive meta-analysis version 3.* Englewood, NJ: Biostat. Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods, 8*, 5-18. https://doi.org/10.1002/jrsm.1230

Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en Ciencias Sociales y de la Salud [Meta-analysis in Social and Health Sciences].* Madrid: Síntesis.

Card, N. A. (2012). *Applied meta-analysis for social science research.* New York, NY: Guilford Press.

Chen, G., Cheng, W., Chang, T., Zheng, X., & Huang, R. (2014). A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education, 1*, 213-225. https://doi.org/10.1007/s40692-014-0012-z

Childwise (2017). *The monitor trends report 2017.* http://www.childwise.co.uk/reports.html#trendsreport.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Courage, M. L. (2017). Screen media and the youngest viewers: Implications for attention and learning. In F. C. Blumberg, & P. C. Brooks (Eds.). *Cognitive development in digital contexts* (pp. 3-28). London: Academic Press.

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*. https://doi.org/10.1080/001401392089673941297-1236

Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2015). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology, 107*, 209-238. https://doi.org/10.1111/bjop.12134.

Duran, E. (2013). Efficiency in reading comprehension: A comparison of students' competency in reading printed and digital texts. *Educational Research and Reviews, 8*(6), 258-269. https://doi.org/10.5897/ERR12.180.

Eyre, J., Berg, M., Mazengarb, J., & Lawes, E. (2017). *Mode equivalency in PAT: Reading comprehension.* Wellington: NZCER. http://www.nzcer.org.nz/system/files/PAT%20Modes_report.pdf.

Gerjets, P., & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educational Psychologist, 38*, 33-41. https://doi.org/10.1207/S15326985EP3801_5.

Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science, 3*, 371-398. https://doi.org/10.1111/j.1756-8765.2010.01081.x

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York, NY: Academic Press. Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions.* The Cochrane Collaboration: Version 5.1.0. [updated March 2011]. http://www.cochrane-handbook.org

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *Journal of Technology, Learning and Assessment, 3*(4). https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1657

Hongler, K. I. (2015). *Superiority of paper as text presentation medium for effective and efficient learning – is it just an illusion?* Master thesis. http://www.mmi-basel.ch/wp- content/uploads/2016/04/hongler_masterarbeit.pdf

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: *Q* statistic or I^2 index? *Psychological Methods, 11*, 193-206. https://doi.org/10.1037/1082-989X.11.2.193

Kingston, N. M. (2008). Comparability of computer- and paper- administered multiple-choice tests for K–12 populations: A synthesis. *Applied Measurement in Education, 22*, 22-37. https://doi.org/10.1080/08957340802558326

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*, 2693-2710. https://doi.org/10.1002/sim.1482

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education, 123*, 138-149. https://doi.org/10.1016/j.compedu.2018.05.005

Koren, D., Seidman, L. J., Goldsmith, M., & Harvey, P. D. (2006). Real-world cognitive—and metacognitive—dysfunction in schizophrenia: A new approach for measuring (and remediating) more "right stuff". *Schizophrenia Bulletin, 32*, 310-326. https://doi.org/10.1093/schbul/sbj035

Kurata, K., Ishita, E., Miyata, Y., & Minami, Y. (2017). Print or digital? Reading behavior and preferences in Japan. *Journal of the Association for Information Science and Technology, 68*, 884-894. https://doi.org/10.1002/asi.23712

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse Processes, 54*, 427-445. https://doi.org/10.1080/0163853X.2017.1319653

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Luyten, H., Merrell, C., & Tymms, P. (2017). The contribution of schooling to learning gains of pupils in Years 1 to 6. *School Effectiveness and School Improvement, Vol. 28*, 374-405. https://doi.org/10.1080/09243453.2017.1297312

Mangen, A., & Kuiken, D. (2014). Lost in an iPad: Narrative engagement on paper and tablet. *Scientific Study of Literature, 4*, 150-177. https://doi.org/10.1075/ssol.4.2.02man

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Margolin, S. J., Driscoll, C., Toland, M. J., & Kegler, J. L. (2013). E-readers, computer screens, or paper: Does reading comprehension change across media platforms? *Applied Cognitive Psychology, 27*, 512-519. https://doi.org/10.1002/acp.2930

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation, 51*, 297-384. https://doi.org/10.1016/S0079-7421(09)51009-2

Mizrachi, D. (2015). Undergraduates' academic reading format preferences and behaviors. *The Journal of Academic Librarianship, 41*, 301-311. https://doi.org/10.1016/j.acalib.2015.03.009

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology, 53*, 17-29. https://doi.org/10.1348/000711000159150

Nishizaki, D. M. (2015). *The effects of tablets on learning: Does studying from a tablet computer affect student learning differently across educational levels.* Senior thesis. http://scholarship.claremont.edu/cmc_theses/1011/

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics, 51*, 1352-1375. https://doi.org/10.1080/00140130802170387

Pansky, A., Goldsmith, M., Koriat, A., & Pearlman-Avnion, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21*, 303-329. https://doi.org/10.1080/09541440802281183.

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences, 26*, 89-102. https://doi.org/10.1016/j.lindif.2013.04.008.

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *The Journal of Technology, Learning, and Assessment, 2*. n6 https://eric.ed.gov/?id=EJ905028.

Pomplun, M., Frey, S., & Becker, D. F. (2002). The score equivalence of paper-and-pencil and computerized versions of a speeded test of reading comprehension. *Educational and Psychological Measurement, 62*, 337-354. https://doi.org/10.1177/0013164402062002009.

Puhan, G., Boughton, K. A., & Kim, S. (2005). *Evaluating the comparability of paper-and-pencil and computerized versions of a large-scale certification test.* ETS Research Report Series, 2005. https://doi.org/10.1002/j.2333-8504.2005.tb01998..x

Rasmusson, M. (2015). Reading paper – reading screen. *Nordic Studies in Education, 35*, 3. http://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva-23417.

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.). *The handbook of research synthesis and meta-analysis* (pp. 295-316). New York, NY: Russell Sage.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638-641. https://doi.org/10.1037/0033-2909.86.3.638.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

Salmerón, L., García, A., & Vidal-Abarca, E. (2018). The development of adolescents' comprehension-based Internet reading skills. *Learning and Individual Differences, 61*, 31-39. https://doi.org/10.1016/j.lindif.2017.11.006.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities, 48*, 369-390. https://doi.org/10.1177/0022219413504995.

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction, 51*, 61-73. https://doi.org/10.1016/j.learninstruc.2017.01.002

Simian, M., Malbrán, C., Fonseca, L., Grinberg, S., Gattas, M. J., Tascón, R., et al. (2016). Reading narrative, expository and discontinuous texts on paper versus screen: Impact on reading comprehension in Argentine high school kids. *Poster presented at 33rd annual meeting of the society for the scientific studies of reading (SSSR), Porto, Portugal.*

Singer, L. M., & Alexander, P. A. (2017a). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education, 85*, 155-172. https://doi.org/10.1080/00220973.2016.1143794

Singer, L. M., & Alexander, P. A. (2017b). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*, 1007-1041. https://doi.org/10.3102/0034654317722961

Thiede, K. W., & 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*, 1024-1037.

https://doi.org/10.1037/0278-7393.25.4.1024

Viechtbauer, W. (2010). *metafor: Meta-Analysis Package for R. R package v. 1.4-0*.

Wang, S., Jiao, H., Young, M. J., Brooks, T., & Olson, O. (2007). Comparability of computer-based and paper-and-pencil testing in K–12 reading assessments. A meta-analysis of testing mode effects. *Educational and Psychological Measurement, 68*, 5-24. https://doi.org/10.1177/0013164407305592

Wolf, M., & Barzillai, M. (2009). The importance of deep reading. *Educational Leadership, 6*, 32-37. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.7284&rep=rep1&type=pdf

Chapter 4

Study 2

Cognitive engagement during print and tablet reading: An eye-tracking study*

Pablo Delgado and Ladislao Salmerón

~

"Thus, reading appears to the casual introspection of the reader. We find, however, that underneath this apparent simplicity, there is an astounding complexity of processes. These have been built up slowly, and by an immense amount of practice, until they have organized and settled into the smoothly running machinery of our present-day reading."

Edmund B. Huey, *The psychology and pedagogy of reading* (1908)

* The content of this chapter is a manuscript in preparation to be submitted for publication.

Chapter 4

Cognitive engagement during print and tablet reading: An eye-tracking study

The present study examined the effect of the interaction between the reading medium and the reading time-frame on cognitive engagement and reading comprehension. One hundred and sixteen undergraduate students read three texts on paper and three texts on a tablet while their eye movements were recorded. Some of them self-paced their reading time, whereas others read under time constraints. After each text, participants completed six questions about the text's content. Before accessing the questions, participants estimated how many of them they expected to answer correctly. Participants' predictions of their performance and their actual performance on the comprehension questions were compared to examine their use of metacognitive monitoring skills. Results showed longer fixation times on text titles and a higher number of fixations on text titles and text sentences when reading on paper than when reading on the tablet. Furthermore, participants' monitoring of their level of comprehension was more accurate at group level when reading on paper because their predictions correlated with their actual performance only in this medium. Finally, reading comprehension scores were slightly higher when reading on paper, although the difference was not significant. Overall, results indicated that participants engaged with the reading task to a greater extent when reading on paper than when reading on a tablet.

4.1. Introduction

Since Prensky (2001) introduced the term *digital natives* to refer to the first generation of students who grew up with digital technologies, there has been an intense debate about whether this new type of learner should be taught pursuant to their learning preferences. According to Prensky's suggestions, they would prefer to learn by reading on a digital medium. However, almost two decades later, the majority of students still opt for printed texts. An international survey covering 19 countries and more than 10,000 university students found that 78% of the participants preferred a print format for academic reading (Mizrachi, Salaz, Kurbanoglu, Boustany, & ARFIS Research Group, 2018). In addition, students expressed concerns about their difficulty in cognitively engaging with the task when reading on a screen, which, according to them, leads to poorer text comprehension and lower memorization. Nevertheless, in contrast to the strong evidence from students' perceptions, there is a lack of literature exploring the difficulties in engaging with a screen during actual reading episodes. To corroborate these effects, in the current study we aimed to experimentally investigate the effect of the reading medium (screen vs. paper) on readers' cognitive engagement by measuring readers' eye-movement behaviour and metacognitive monitoring, and on reading comprehension outcomes.

4.1.1. Reading medium and text comprehension

Reading on screens is ubiquitous in today's society, and it is increasingly present in instructional practices and educational policies. For example, almost 40% of the 4^{th} grade students participating in the international evaluation program of student learning in reading, PIRLS 2016, spend more than 30 minutes a day finding and reading information on the Internet (Mullis, Martin, Foy, & Hopper, 2016). Similar results on the frequency of use of computers for learning purposes were found by the PISA international program evaluating Secondary students in the OECD countries (OECD, 2015). Nevertheless, strong concerns have been raised about the possible harmful effects of the pervasive use of digital technologies on learning (e.g., Greenfield, 2015) and, more specifically, on reading (e.g., Baron, 2015; Wolf, 2018). Within this research field, scholars have been examining whether comprehension outcomes are equivalent across reading media. Findings are heterogeneous, but the majority show better comprehension for on-print reading when media differences appear (e.g., Ackerman & Lauterman, 2012; Lenhard, Schroeders, & Lenhard, 2017; Mangen, Walgermo, & Brønnick, 2013; Singer & Alexander, 2017b; Wästlund, Reinikka, Norlander, & Archer, 2005), at least under some circumstances –e.g., the type of text, task, or even in some readers (Singer & Alexander, 2017a)–.

Indeed, the results from three recent meta-analyses yielded better comprehension for on-print reading in both cases. The mean effect sizes ranged from a Hedges' g of −0.21 (Delgado, Vargas, Ackerman, & Salmerón, 2018; Kong, Seo, & Zhai, 2018) to −0.25 (Clinton, 2019), and this on-screen inferiority was greater when reading expository texts, $g = -0.32$ (Clinton, 2019) and $g = -0.27$ (Delgado et al, 2018). Although these effect sizes would be considered small according to Cohen's benchmarks (1988), they are located within Hattie's (2009) zone of *Teacher effects*, which represents "what teachers can accomplish in a typical year of schooling" (p. 20). Additionally, Delgado et al. (2018) also found a larger effect size in studies that imposed reading time constraints, $g = -0.26$, vs. those in which participants self-paced their reading time, $g = -0.09$. However, the majority of the studies comparing on-print reading with on-screen reading have used either desktop or laptop computers; only recently, a few studies have started to analyse the effects of tablets (Delgado et al., 2018). Because the postural and physical experience of using tablets is closer to that of reading print than that of computers, we cannot assume that the detrimental effects of computers would be equivalent when reading with tablets (e.g., Hermena et al., 2017; Sackstein, Spark, & Jenkins, 2015). Our study aims to add to this growing number of studies by comparing students' reading of tablets and print documents while keeping their body posture constant.

According to the Shallowing hypothesis, on-screen inferiority is due to a tendency to process information from screens superficially (Annisette & Lafreniere, 2017; Delgado et al., 2018; Wolf & Barzillai, 2009). From this perspective, reading on screens prompts low cognitive engagement due to the habits people develop in their interactions with digital media, which are frequently characterized by quick and superficial interactions with the information (Baron, 2015). Kaufman and Flanagan (2016) experimentally tested this assumption in a series of studies with young adults. They found higher scores on inferential comprehension questions in participants who read the printed version of a text, whereas participants in the digital condition showed better memory for specific details (Study 2). They also found better results in the participants who completed a problem-solving task that required using high-level processing strategies on paper (Study 3A). Finally, they replicated the latter task only on a screen and splitting the sample into three groups. Two groups were primed to activate a high or low construal cognitive level, respectively, whereas the third group was not primed. The results indicated that the low-level-primed group and the non-primed group performed similarly, whereas the high-level-primed group outperformed them (Study 3B). The authors argued that the digital medium, per se, triggers a low level of cognitive engagement that results in a shallow processing mind-set, as the Shallowing Hypothesis proposes. Current evidence for the Shallowing hypothesis relies heavily on product measures, such as responses on a text comprehension questionnaire, but there is almost no evidence from studies analysing differences in cognitive engagement based on the medium during the reading process.

Our study aims to fill this gap.

4.1.2. Eye movements as an indicator of cognitive engagement during reading.

Cognitive engagement is a core learning process that can be defined as investing the necessary mental effort to comprehend complex ideas and master difficult skills (see Friedicks, Blumenfeld, & Paris, 2004). Visual attention, as measured by the eye-tracking technique, has been found to be the most valuable indicator of cognitive engagement (D'Mello, Dieterle, & Duckworth, 2017; Miller, 2015). According to the *mind-eye hypothesis*, textual information is processed during the time that eyes stay fixated on words (Just & Carpenter, 1980). Although fixation times on words tend to be quite homogeneous, variations in fixation times are considered an indicator of different levels of engagement with reading. Thus, longer fixations are generally associated with deeper processing (Holmqvist et al., 2011). For example, studies on text relevance have found longer fixations when reading text passages that were relevant to the task goals (Kaakinen, Hyönä, & Keenan, 2002; Kaakinen & Hyönä, 2005, Kaakinen, Ballenghein, Tissier, and Baccino; 2018). Moreover, comprehending difficult texts, passages, or sentences within a text, which requires more effortful processing, results not only in longer fixation times, but also in shorter saccades and more regressions[5] (Kinnunen & Vauras, 1995; Rayner, Chace, Slattery, & Ashby, 2006; Rayner, 1998).

Fixations on sentences already read (so-called *look-back* fixations or *second-pass* fixations) are specifically considered indicators of further efforts to comprehend and memorize the text (Hyöna, Lorch, & Kaakinen, 2004; Hyönä & Nurminen, 2006; Rayner

[5] Saccades are usually defined as the rapid eye movements between fixations when we read, look at a scene, or search for an object (Rayner, 1998, p. 373). Regressions are specific saccades that occur when the reader's eyes do right-to-left movements along the line or movements back to previously read lines (p. 375).

et al., 2006). In this regard, Ariasi and Mason (2011) measured content comprehension and fixation times in undergraduate students when reading a refutational expository[6] text, compared to a standard non-refutational text. Results indicated that those who read the refutational text made longer fixations, particularly when rereading some parts of the text, and so their reading outcomes revealed deeper comprehension, as indicated by their conceptual knowledge gain, measured by pre-test and post-test questions about the text's content. Longer second-pass fixation times have also been related to high scores on a self-reported inventory on learning strategies (Catrysse et al., 2018) and to the use of metacognitive monitoring abilities (Kinnunen & Vauras, 1995; van Gog & Jardozka, 2013).

Eye movements have also been used to identify how readers process structural elements of the text, such as headings, an indicator of strategic reading. For example, Hyönä, Lorch, and Kaakinen (2002) identified four different profiles of adult readers, and those they classified as *topic structure processors* were characterized by spending much longer fixation time on headings. Therefore, this type of reader wrote better summaries of the texts. Moreover, the presence (thus, the reading) of headings has been found to speed up the processing of subsequent sentences, especially the initial sentence introducing the topic of the text (Hyöna & Lorch, 2004), as well as facilitating readers' construction of the structure of the text (Lorch, Pugzles, Lorch, McGovern, & Coleman, 2001) and fostering information recall, especially of unfamiliar topics (e.g., Lorch & Lorch, 1996).

In sum, the research corpus on eye movements while reading shows that cognitive engagement, whether induced by the text (e.g., text difficulty), the task (e.g., reading goals), or the reader him/herself (e.g., the use of reading strategies), is usually reflected

[6] Refutational texts directly refute the reader's prior misconceptions about the topic at hand, fostering conceptual change (Hynd, Whorter, Phares, & Suttles, 1994).

in longer total fixation times, especially when re-visiting text segments already read. Based on this, if on-screen reading hinders cognitive engagement, as predicted by the Shallowing Hypothesis, we would expect shorter fixation times and/or a lower number of fixations when reading on this medium, for both first-pass reading and, especially, second-pass fixations.

To our knowledge, only Bansi et al. (2016), Kretzschmar et al. (2013), and Latini, Bråten and Salmerón (2019) have compared eye movements and reading comprehension across reading media within the same experiment. Latini et al. (2019) recorded university students' eye movements while reading an illustrated text in order to write a report based on the text content. It was necessary to integrate the textual content and the illustration in order to gain a good understanding of the information, which was measured by scoring the reports based on the number of relevant ideas included from the document and adding extra points for ideas that combined information from the text and the picture. Results indicated that participants who read the document on paper engaged in in-depth reading to a greater extent than those who read it on a computer, as indicated by an increased integrative processing of the textual and pictorial information (i.e., a higher number of gaze transitions between the text and the picture). Accordingly, the on-print readers included more integrative ideas from the document. The authors argued that these results supported the assumptions of the Shallowing hypothesis in a multimedia learning context. Nevertheless, it should be noted that the authors did not compare other eye-movement parameters such as fixation times on the text.

However, Bansi et al. (2016) and Kretzschmar et al. (2013) took a different approach by examining fixation times on texts without pictures. Bansi et al. (2016) asked a sample of undergraduate students to read three texts on paper and three texts on a screen (187-234 words). Given that they used a screen-based eye-tracking, the printed texts were stuck on the screen. After reading on each medium, the participants answered five questions per text (15 questions per medium), including both literal and inferential questions. The eye-movement measures were the total number of fixations, the forward fixation duration, and the forward saccade length. Their results showed longer forward fixations when participants read in print, and longer forward saccades when reading on the screen. Given the lack of differences in comprehension outcomes, the authors proposed three possible interpretations for the differences in the eye-tracking measures: 1) on-print reading is more effortful; 2) on-print reading leads to more reading engagement; or 3) the readers apply different strategies on each reading media.

Kretzschmar et al. (2013) conducted their study in two different samples of young and elderly adults (mean age: 25.7 and 66.8 years), respectively. Reading media effects on eye movements and comprehension, as well as on EEG activity, were compared within participants by means of nine short texts (176−266 words, three texts per medium) and a sentence verification task consisting of two questions per text about details from the content (e.g., "Years ago, Birte forgot to close the lion cage."; which required a 'no' answer because a correct sentence would have been 'Birte forgot to close the lama cage'.). Reading time was self-paced by participants. Findings showed that, although young adults' eye movements did not differ across media, older adults showed longer fixation times when reading printed texts (vs. on a tablet). Given that the results yielded no differences in performance across media, the authors suggested that the increased total fixation time in the older adults when reading on paper was due to difficulties related to text-background discrimination. The tablet screen could provide increased contrast due to backlighting, and so, according to the authors, it helped older readers to read with less effort. Other interpretations are possible, however, because the lack of effects on the reading task could be due to the study characteristics.

Some methodological features of Bansi et al.'s (2016) and Kretzschmar et al.'s (2013) studies limited the possibility of providing a robust test for the Shallowing hypothesis. Besides both studies used short texts, on the one hand, Bansi et al. stuck the printed texts on the eye-tracker screen, so the reading situation was similar to that of reading digital texts. On the other hand, Kretzschmar et al. used comprehension questions covering only literal comprehension, and they only utilized one eye-movement measured (summed fixation time per page). Thus, we aim to build on these two studies by performing a more robust test of the Shallowing hypothesis by including more challenging reading situations, such as reading longer texts, reading under time pressure (Walczyk, 1995), and more complex questions that test both literal and inferential comprehension –i.e., connections between propositions not explicitly connected in the text (Kintsch & van Dijk, 1978)–, which require a higher level of cognitive operations (Kintsch, 1998). Additionally, we shall increase the diversity of eye-tracking measures in order to discriminate first-pass and second-pass reading times and the number of fixations at the sentence level.

4.1.3. Metacognitive monitoring deficits as the cause of the on-screen inferiority.

Another indicator of cognitive engagement is the use of metacognitive abilities, such as monitoring, because they provide the basis for the allocation of effort, cognitive resources, and the use of learning strategies (Fiedler, Ackerman, & Scarampi, 2019; Hacker, Bol, & Keener, 2008). Metacognitive monitoring, or the on-line evaluation of one's comprehension and text relevance, is the basis for regulatory decisions that involve, for instance, rereading a previous paragraph that has been perceived as important. In-depth reading relies so much on monitoring that deficits in this skill could provide an explanation for shallower reading when reading on screen (e.g., Ackerman & Lauterman, 2012). Inefficient monitoring would misguide readers in their allocation of effort and strategies, which would potentially result in lower comprehension if the task is challenging enough, such as in situations where readers have limited time to read (Delgado et al., 2018).

Metacognitive *calibration* is the most widely studied monitoring skill in research on the effects of the reading medium. It can be defined as 'the degree to which a person's perception of performance corresponds with his or her actual performance' (Hacker et al., 2008, p. 433). It represents the absolute accuracy between the a priori estimation of the performance level on a test and the actual performance. Calibration can be measured by asking participants to indicate how many items they expect to answer correctly out of the total number of items on a test (Hacker et al., 2008). A complementary measure of metacognitive-monitoring is *resolution*, which represents the relative accuracy of the metacognitive judgements. Namely, resolution is the extent to which people's judgments discriminate between higher and lower levels of performance (Rawson, Dunlosky, & Thiede, 2000; Thiede, Anderson, & Therriault, 2003). Thus, resolution is the extent to which predictions are correlated with the actual performance. Therefore, people may calibrate poorly but show good resolution.

Research on the effect of the reading medium on readers' metacognitive monitoring has also yielded complex patterns. Clinton's (2019) meta-analysis reported an overall mean effect of medium, with readers' calibration found to be better when reading on paper than when reading on screen ($g = 0.20$). However, this result is based on a relatively small sample of effect sizes ($k = 11$) provided by only six primary studies. Moreover, when looking at these studies individually, the results are highly diverse and inconclusive. On the one hand, two studies by Ackerman et al. (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012) and a study by Singer Trakhman, Alexander, and

Berkowitz (2019) reported that participants calibrated significantly worse when reading on screen. However, other studies found that participants' calibration did not significantly differ across reading media (Chen & Catrambone, 2015; Delgado & Salmerón, 2019; Norman & Furnes, 2016). Additionally, Lauterman and Ackerman (2014) found better calibration, as well as an improvement with practice[7], in participants who read on their preferred medium, whether on paper or on screen.

Overall, in spite of the heterogeneity of the results presented above, shallower on-screen reading could be explained, at least in part, by a deficit in readers' metacognitive monitoring when reading texts in this medium. Thus, we aim to shed more light on this issue by comparing participants' metacognitive monitoring when reading on screens vs. printed texts under two different reading time-frame conditions (i.e., self-paced vs. pressured time) because this latter factor influences the effect of the reading medium on both calibration and reading comprehension (Ackerman & Lauterman, 2012; Delgado et al., 2018).

4.1.4. Overview of the present study

Extending the research corpus on the effect of reading on tablets, a digital device that is still under-studied, the main goal of our study was three-fold: 1) to test the on-screen cognitive disengagement proposed in the Shallowing hypothesis by comparing readers' eye movements when reading on tablets and when reading printed texts; 2) to measure participants' metacognitive monitoring while reading on both media and under two different reading time-frame conditions (self-paced vs. pressured time); and 3) to compare reading comprehension outcomes across the two reading media.

[7] Participants in this study read three texts in a row. They made a prediction of performance and then completed a test after each text, so that their calibration was measured for each text.

For the aforementioned purposes, participants in our study read three texts on each medium while their eye movements were tracked. Moreover, a group of participants self-paced their reading time, whereas others read under time pressure (Ackerman & Lauterman, 2012), a manipulation that we expected to affect reading performance, yielding a larger medium effect under time pressure (Delgado et al, 2018). We measured eye movements, text comprehension, and metacognitive calibration. Additionally, we controlled for confounding variables by measuring participants' reading medium preference (Lauterman & Ackerman, 2014), prior knowledge (Kendeou & van den Broek, 2007), and interest in each text topic, both prior interest and situational interest (Schraw & Lehman, 2001). The questions that guided our study were:

- Q1: Does participants' cognitive engagement, as measured by eye fixation times and number of fixations, differ across reading media and reading time-frame conditions?
- Q2: Does participants' metacognitive calibration differ across reading media and reading time-frame conditions?
- Q3: Does reading comprehension differ across reading media and time-frame conditions?

4.2. Method

4.2.1. Participants

One hundred and sixteen undergraduate students from the Teacher education or Psychology degrees at a large Spanish university participated in the study in exchange for course credit or a small monetary compensation (mean age = 20.41 years, *SD* = 2.68; 84.49% female). Participants' eye movements while reading were recorded. Unfortunately, eye-tracking data for 34 participants were lost for unforeseen technical reasons after the completion of the study (the hard drive containing the data files was broken). In addition, eye-tracking data from 11 other participants were not used due to inaccurate calibration. Because these factors did not interfere with the participants' actual performance during the study, we decided to keep the original sample of 116 participants for the analyses that only involved calibration and comprehension measures, and we used a subsample of 71 participants for the eye-tracking analyses. All participants signed an informed consent, and none of them had any learning difficulties.

We conducted a priori G-power analyses (G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2017) based on the interaction effect of the medium and reading time-frame on comprehension ($\eta^2_p = .07$) and calibration ($\eta^2_p = .06$), reported by Ackerman & Lauterman (2012), and on the *t* statistic reported by Rayner et al. (2009, Experiment 1) for differences in fixation times ($t = 4.16$) and number of fixations ($t = 6.90$) between reading difficult and easy passages. Results indicated that a sample of 44 participants would be necessary to detect such effects for a two-factor mixed design ($\alpha = .05$, $\beta = .05$). Therefore, our sample size was appropriate for our purposes.

4.2.2. Apparatus

Eye-tracker. Participants' eye movements were recorded with the SMI Eye-Tracking Glasses version 1, a video-based, infrared glasses-type eye tracker with a binocular sampling rate of 30Hz and automatic parallax correction. Using an eye-tracking glasses system allowed us to record participants' eye movements when reading on paper. This eye-tracker has an accuracy of 0.5º over all distances and a gaze tracking range of 80º horizontal and 60º vertical. System calibration was performed using a 3-point scheme. To maximize data accuracy, the system was calibrated 6 times per participant over the whole experiment (before starting to read each text).

Reading devices. Participants read the digital version of the texts on a 9.7-inch tablet (SPC Glow 9.7) with a 1024 x 768 screen resolution and brightness set at 50%. Participants turned the text pages by swiping the screen with a finger or tapping the side of the screen. Printed texts were presented in a ring binder whose pages mimicked exactly the size (approximately DIN A5 size) and formatting of the digital version (see Fig. 4.1). The booklet was printed on one side, so that only one page was shown at a time (which also mimics the text on the tablet). Both devices were placed on a stand located approximately 60−70 cm from the participants, who were not allowed to hold them. This was done to minimize the offset of eye movements during reading.

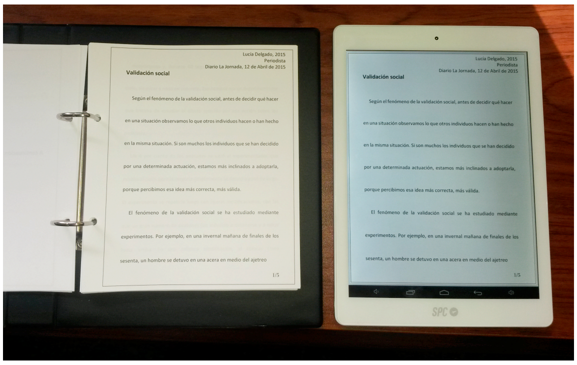

Figure 4.1. On-print and on-screen versions of the same page of one of the texts.

4.2.3. Materials

Texts. We used six expository texts[8] adapted from the Spanish version of Scientific American magazine, ranging from 362 to 457 words (4 to 6 pages, approximately A5 size). According to the Inflesz Scale for Spanish texts (Barrio-Cantalejo et al., 2008), which is based on word length and number of syllables, one text showed a readability index of 56.15 (*Normal difficulty:* e.g., newspapers, magazines), four texts ranged from

[8] See Appendix B

54.96 to 40.27 (*Somewhat difficult*: e.g., scientific-dissemination texts, specialized press), and one text scored 38.00 (*Very difficult:* e.g., undergraduate textbook, scientific text). The font was 12-point Calibri, and lines were triple-spaced in order to enhance eye-tracker accuracy.

4.2.4. Dependent variables

Eye fixation times. Three eye-movement measures were calculated at the sentence level by averaging fixations and fixation times for each text title and text sentence in the 6 texts: 1) *Fixation time per word* (in *ms*), 2) *Fixations per word*, and 3) *Regressive fixations per word* (i.e., number of fixations that occur after a saccadic movement from right to left before exiting the sentence currently read; Rayner, 1998). Each measure was computed separately for *First-pass* reading (i.e., eye movements during the first reading of a sentence) and *Second-pass* reading (i.e., eye movements on a sentence that had already been fully read). First-pass reading measures were calculated by including first-pass fixations and additional fixations on each sentence when readers left the sentence before reading it completely and returned to it before fixating later sentences (Hyöna, Lorch, Rink, 2003). Additionally, *Second-pass reading time* (i.e., total time, in *ms*, dedicated to second-pass reading) was also calculated for each participant. A summary of these measures can be found in Table 4.1. Measures 1 to 6 (see Table 4.1) were calculated separately for the text titles and the text sentences.

Metacognitive calibration and resolution. After reading each text, participants predicted their performance on the comprehension questions by indicating how many questions, out of the total of six, they estimated that they would answer correctly. Predictions of performance (POP) were compared to the actual performance within each group as a measure of calibration –see Ackerman & Lauterman (2012) for a similar procedure–.

Table 4.1. *Description of each eye-tracking measure*

Eye-tracking measure	**Description**
1. *First-pass fixation time per word*	Time that eyes kept fixated during the first reading of a text title or a text sentence divided by the number of words in the sentence or the title.
2. *Second-pass fixation time per word*	Time that eyes kept fixated when rereading a title or a sentence that has been completely read before, divided by the number of words in the sentence or the title.
3. *First-pass fixations per word*	Number of fixations during the first reading of a title or a sentence, divided by the number of words in the sentence or the title.
4. *Second-pass fixations per word*	Number of fixations when rereading a title or a sentence that has been completely read, divided by the number of words in the sentence or the title.
5. *First-pass regressive fixations per word*	Number of fixations that occurs after a saccadic movement from right to left during the first reading of a title or a sentence, divided by the number of words in the sentence or the title.
6. *Second-pass regressive fixations per word*	Number of fixations that occurs after a saccadic movement from right to left during the rereading of a title or a sentence that has been completely read before, divided by the number of words in the sentence or the title.
7. Rereading time	Total summed fixation time per text.

Additionally, a *Calibration index* for each participant and each text was computed by subtracting the number of correct answers from his/her POP. These scores were added together per medium, so that we computed a *Calibration index* for each reading medium per participant. Negative values indicated underestimation of the actual performance on the comprehension questions, and positive values indicated overconfidence.

For metacognitive resolution, the usual approach to metacognitive resolution is based on examining the correlation between judgements and performance across trials within each participant (e.g., Metcalfe & Finn, 2008). However, given that we had only 3 measures of POPs per participant and medium (according to the comprehension questions, see below), we adapted this approach to group level by examining the correlation between POPs and comprehension scores for each medium condition.

Reading comprehension. Participants' comprehension of the texts' content was measured with 6 multiple-choice questions (four alternatives) per text, half covering literal comprehension (i.e., questions about specific details from a single sentence) and half covering inferential comprehension (i.e., questions requiring them to relate two ideas not explicitly connected in the text)[9]. In all, the participants answered 36 questions, 18 per reading medium. Given that each participant in our study completed half of the questions in each medium, the reliability of this measure was examined based on the results from a pilot study in which 24 participants read all the texts and answered all the questions on a desktop computer. Due to the nature of the items on this task, we used the omega coefficient (McDonald, 1999) to test the reliability because it is considered a better alternative to Cronbach's alpha for non-ordinal items with less than 5 response options (Viladrich, Angulo-brunet, & Doval, 2017). Results indicated good reliability for the 36 questions ($\omega = .88$).

4.2.5. Covariates

Prior knowledge and interest in text topics. Similar to Singer and Alexander (2017b), participants rated their perceived prior knowledge about each text topic from 1 (*I know nothing*) to 10 (*I am an expert*). Previous research has found perceived knowledge to be a valid measure (Stanovich & West, 2008). Participants also rated their interest in each test topic from 1 (*Not interested at all*) to 10 (*Very interested*), following the procedure by Fulmer and Frijters (2011). An average score on each measure was calculated per medium and participant, ranging from 1 to 10.

Situational interest. Once participants had completed the reading task, they rated their situational interest in each text from 1 (*Not interested at all*) to 10 (*Very interested*) by

[9] See Appendix B

answering the question "How interested were you in the topic of the text?" (adapted from Unsworth & McMillan, 2013). Similarly, an average score was calculated per medium for each participant.

Medium preference and use. Participants reported their preferred medium for learning purposes by completing a 5-item questionnaire (two reversed items). Each item presented a statement about the use of printed or digital texts, rated from 1 *(I strongly disagree)* to 5 *(I strongly agree)* (e.g., *If I have notes from any course in a computer file, I prefer to read them on the computer or any other digital device rather than printing them*). An average rating score was computed for each participant, ranging from 1 (strong preference for reading on paper) to 5 (strong preference for reading on digital device). Reliability for this measure was good (ω = .86). In addition, participants completed information about their daily use (in hours) of digital technologies for leisure and learning/professional purposes.[10]

4.2.6. Procedure

Participants were tested in a quiet room in a single individual session that lasted approximately 75 minutes. After completing the prior knowledge and topic interest questionnaire, participants were equipped with the eye tracking glasses. Then, each participant read the experimental instructions on the same reading device he/she used to start the experimental task: "*You are now going to read six short texts, and you will answer six multiple-choice comprehension questions after each text. You won't be allowed to go back to the text while answering the questions. You will be asked to predict how many questions you will answer correctly before knowing what they are. You will first read a practice text and answer only four questions in order to familiarize yourself*

[10] See Appendix B

with the task. You can read the texts at your own pace. After you finish each text, you will be required to make your prediction about performance. You will find the questions on the following pages". Participants who completed the task under time pressure had 70% of the mean time that participants in the self-paced time group took to read each text (Ackerman & Lauterman, 2012). They received the following instruction: "*For each text you will have only 70% of the time that a group of people took on average when reading the same text at their own pace. You have to be efficient. After the time is over, you will be required to turn the pages to complete the performance prediction and answer the questions*". Each participant read the first three texts (plus the practice text) on the same medium, and the remaining three texts on the other medium. The order of the texts was randomized for each participant, and the order of the reading media was counterbalanced between participants. Participants answered the questions for each text in the same medium they used to read the text. Finally, they filled out the questionnaires on situational interest in texts, medium preference, and the use of digital devices.

4.3. Results

We first explored the data distribution for each dependent variable (i.e., reading comprehension score, POPs, calibration index, and eye-tracking measures) to identify outlier scores (±2 *SD from the mean*), which were removed from subsequent analyses. In addition, data on participants' medium preference, situational interest in both media, and eye movements were normalized following the two-step approach proposed by Templeton (2011). In the first step, each variable was transformed into a percentile rank, yielding uniformly distributed probabilities. In the second step, the latter results were transformed to normally distributed z-scores by applying the inverse-normal transformation.

4.3.1. Covariates

Outliers for each covariate (±2 *SD from sample mean*), representing less than 5% of the data in all cases, were also removed and replaced by the sample mean. Descriptive data for the covariates can be found in Table 4.2. There were no differences between the reading time-frame groups on any of these variables (all *F*s < 2.93 and all *p*s > .09. There were no significant correlations between the covariates and the scores on the comprehension questions and the calibration index in any case. However, some covariates were significantly related to the participants' POPs and some of the eye-movement measures (see Table 4.3 and Table 4.4).

In order to control for confounding effects, we included all the covariates described above in subsequent ANCOVAS when examining differences in the dependent variables across the experimental conditions. Following the recommendations of Schneider, Avivi-Reich, and Mozuraitis (2015) for within-participants factors, the covariates were centered prior to entering them into the analyses.

Table 4.2. *Descriptive data for the covariates*

	N	***M***	***SD***	**Skewness**	**Kurtosis**
Use of digital devices for prof./study purposes	116	3.89	1.95	0.38	-0.81
Use of digital devices for leisure	116	3.33	1.76	0.44	-0.64
Medium preference	116	1.91	0.66	0.71	-0.12
Prior knowledge on printed texts topic	116	3.32	1.43	0.26	-0.80
Prior knowledge on digital texts topic	116	3.29	1.42	0.12	-0.86
Prior interest on printed texts topic	116	6.22	1.38	-0.20	-0.70
Prior interest on digital texts topic	116	6.27	1.28	-0.02	-0.43
Situational interest on printed texts content	116	6.46	1.27	-0.01	-0.30
Situational interest on digital texts content	116	6.55	1.32	-0.02	-0.33

Table 4.3. *Person correlations between the study variables in the on-print reading condition*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Comprehension																
2. POP	.25**															
3. Calibration Index	-.68**	.49**														
4. Prior knowledge	.00	.26**	.11													
5. Prior interest	-.08	.24**	.17	.44**												
6. Situational interest	-.07	-.09	-.01	-.13	-.08											
7. Medium preference	.11	.04	-.08	.11	.01	.02										
8. F-pass fixation time on titles	.04	-.10	-.03	-.15	.01	.04	-.13									
9. S-pass fixations on titles on titles	-.11	-.03	.04	-.26*	-.02	.04	-.25*	.86**								
10. F-pass regressive fixations on titles	-.03	.06	.05	-.22	.04	.01	-.37**	.69**	.85**							
11. F-pass fixation time on texts	.03	-.11	-.01	-.14	-.08	.18	-.18	.23	.14	.04						
12. S-pass fixation time on texts	.13	.28*	.05	-.03	.04	.16	-.16	.04	.10	.16	.31*					
13. F-pass fixations per word on texts	-.02	.02	.03	-.18	-.01	.25*	-.11	.06	.16	.14	.75**	.38**				
14. S-pass fixations per word on texts	.14	.18	.03	-.10	.04	.16	.00	-.06	-.07	-.02	.29*	.80**	.45**			
15. F-pass regressive fixations on texts	.15	.19	.04	-.17	-.04	.24*	-.05	.13	.19	.26*	.51**	.36**	.78**	.37**		
16. S-pass regressive fixations on texts	.21	.30*	.10	-.01	.10	.27*	-.07	.12	.10	.19	.32*	.60**	.43**	.69**	.54**	
17. Rereading time	.10	.21	.11	-.02	.08	.04	-.05	.05	-.02	.01	.33**	.65**	.37**	.89**	.30*	.83**

Table 4.4. *Person correlations between the study variables in the on-screen reading condition*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Comprehension																		
2. POP	.02																	
3. Calibration	-.65**	.66**																
4. Prior knowledge	.06	.26**	.08															
5. Prior interest	.05	.23*	.14	.37**														
6. Situational interest	-.11	-.12	-.08	-.05	-.14													
7. Medium preference	-.08	-.09	-.07	.15	.14	.03												
8. Use of digital devices for prof/study	.08	.06	-.01	.06	.18	.03	-.01											
9. Use digital devices for leisure	-.19	.02	.08	-.06	-.06	.01	-.14	-.05										
10. F-pass fixation time on titles	-.01	-.14	-.02	-.21	-.04	.14	-.22	.09	-.04									
11. S-pass fixations on titles on titles	.01	-.16	.00	-.29*	-.06	.12	-.16	.08	-.10	.86**								
12. F-pass regressive fixations on titles	.09	-.08	.04	-.01	.08	-.16	-.03	.07	-.12	.37**	.56**							
13. F-pass fixation time on texts	-.17	-.26*	.03	-.21	-.23	.15	-.23	-.06	.06	.29*	.28*	.12						
14. S-pass fixation time on texts	.36**	.24	-.02	.06	-.12	.09	-.14	.02	.03	-.12	-.12	-.08	.11					
15. F-pass fixations per word on texts	.00	-.17	.02	-.38**	-.19	.15	-.13	-.08	-.08	.17	.27*	-.02	.80**	.21				
16. S-pass fixations per word on texts	.37**	.16	-.02	-.01	-.09	.07	-.22	.02	.02	-.07	-.05	-.04	.18	.82**	.27*			
17. F-pass regressive fixations on texts	.06	-.06	-.03	-.28*	-.04	.05	-.11	-.11	-.05	-.02	.08	.12	.58**	.08	.79**	.23		
18. S-pass regressive fixations on texts	.32*	.15	-.10	.02	-.17	.16	-.15	.01	-.11	-.12	-.11	-.17	.10	.83**	.25	.79**	.16	
19. Rereading time	.38**	.20	-.04	-.02	-.14	.12	-.16	-.09	.00	-.05	-.03	-.06	.17	.74**	.29*	.90**	.22	.86**

Table 4.5. *Descriptive statistics for each variable in each experimental condition. Skewness and Kurtosis for each variable in each medium condition.*

Time-frame condition	***Self-pace reading time***				***Pressured reading time***				**Whole Sample**			
Medium condition	***On-print***		***On-screen***		***On-print***		***On-screen***		***On-print***		***On-screen***	
	n	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)	Skew.	Kurt.	Skew.	Kurt.
Comprehension score	65	8.91 (2.43)	67	8.33 (2.52)	45	7.98 (2.52)	44	7.36 (2.41)	-0.09	-0.47	0.28	-0.55
POPs	68	9.94 (2.36)	66	9.68 (2.21)	44	9.00 (1.75)	46	9.20 (2.41)	0.13	-.48	0.06	-0.64
Calibration Index	68	0.81 (3.09)	65	1.15 (3.06)	46	1.35 (2.7)	46	1.35 (3.14)	-0.15	-0.88	-0.09	-0.49
First-pass fixation time per word on titles	31	271.09 (101.48)	31	215.07 (85.43)	39	261.13 (86.27)	39	237.60 (52.83)	0.00	-0.37	0.00	-0.37
First-pass fixations per word on titles	32	1.66 (0.52)	31	1.30 (0.43)	40	1.58 (0.54)	41	1.40 (0.30)	0.00	-0.37	0.00	-0.36
First-pass regressive fixations per word on titles	32	0.60 (0.20)	31	0.40 (0.18)	38	0.52 (0.20)	39	0.42 (0.15)	0.00	-0.38	0.00	-0.37
First-pass fixation time per word on sentences	29	204.94 (38.45)	29	199.81 (45.20)	41	187.41 (24.40)	41	184.07 (25.24)	0.00	-0.37	0.00	-0.37
Second-pass fixation time per word on sentences	30	92.05 (41.62)	30	86.13 (42.60)	32	61.55 (31.03)	31	59.41 (34.59)	0.00	-0.39	0.00	-0.39
First-pass fixations per word on sentences	30	1.14 (0.20)	30	1.13 (0.22)	41	1.05 (0.13)	41	1.03 (0.13)	0.00	-0.37	0.00	-0.37
Second-pass fixations per word on sentences	29	0.47 (0.18)	29	0.41 (0.20)	41	0.24 (0.17)	41	0.21 (0.21)	0.08	-0.53	0.10	-0.56
First-pass regressive fixations on sentences	31	0.28 (0.09)	31	0.27 (0.10)	41	0.22 (0.07)	41	0.21 (0.07)	0.00	-0.36	0.00	-0.37
Second-pass regressive fixations on sentences	27	0.03 (0.02)	27	0.03 (0.02)	35	0.02 (0.01)	34	0.02 (0.02)	0.00	-0.39	0.11	-0.59
Rereading time	28	9774.5 (5056.7)	30	9743.2 (5961.2)	41	3481.2 (4867.1)	41	2900.6 (6123.5)	0.08	-0.53	0.10	-0.55

4.3.2. Effects of reading medium and time-frame on eye movements.

Eye-movements on text titles.

We conducted a two-way mixed ANOVA for each measure of participants' eye movements when reading the text titles, with the medium (within) and reading time-frame (between) as independent variables. ANCOVA assumptions regarding normality and homoscedasticity were met in all cases. In this set of analyses, we included only data from the participants' first-pass reading because only 12 participants reread this part of the text in both media. Descriptive data for these measures can be found in Table 4.5.

Participants' prior knowledge and medium preference correlated with the number of *First-pass fixations per word* on titles. Medium preference also correlated with the number of *First-pass regressive fixations per word* on titles. Hence, these covariates were controlled in each case. ANOVA results showed a main effect of medium indicating that the *First-pass fixation time per word* on titles was longer when reading on paper than when reading on the tablet, $F(1, 65) = 11.97$, $p < .01$, $\eta^2_p = .16$. In the same direction, ANCOVA results showed that the mean number of *First-pass fixations per word* on titles, $F(1, 67) = 17.09$, $p < .001$, $\eta^2_p = .21$, and the number of *First-pass regressive fixations per word on titles* $F(1, 65) = 25.48$, $p < .001$, $\eta^2_p = .28$, were also higher when reading on paper. Finally, the main effect of time-frame and the interaction effect of medium and time-frame on these three measures of readers' eye movements were not significant (all Fs < 1.30, all ps > .26).

Eye-movements on text sentences.

Means and *SD*s for each of the seven eye-tracking measures when reading text sentences are also presented in Table 4.5. Testing of the ANOVA and ANCOVA assumptions revealed that homoscedasticity was met only for the analyses of *First-pass* and *Second-pass regressive fixations*, and so ANCOVAs were performed. In the remaining cases, we performed robust estimation ANOVA for mixed designs, a non-parametrical alternative to mixed ANOVA based on the analysis of trimmed means (Wilcox, 2012). Given that trimmed means are calculated by removing the tails of the data distribution, outliers were not excluded when we used this method.

Robust mixed ANOVAs were conducted using the WRS2 package for R (Mair & Wilcox, in press), with the trimming level set at 10% in all cases. Results showed no significant effects of medium on the *First-pass fixation time per word,* $F(1, 43.05) = 2.29$, $p = .13$, or on the *Second-pass fixation time per word,* $F < 1$; a significant effect of reading time-frame in both cases, $F(1, 30.19) = 9.55$, $p < .01$, and $F(1, 49.86) = 16.53$, $p < .001$, revealing that participants fixated and refixated longer per word when self-pacing their reading time. No interaction effects were found between medium and time frame (in both cases, $Fs < 1$).

Regarding the number of *First-pass* and *Second-pass fixations per word*, robust mixed ANOVA indicated that the effect of medium was significant for first-pass reading, $F(1, 43.11) = 4.43$, $p = .04$, and non-significant for second-pass reading, $F(1, 46.14) = 3.57$, $p = .06$, indicating that participants fixated, and tended to refixate, more often when reading and rereading on paper than when reading on tablets. Moreover, the effect of the reading time-frame was significant in both cases, $F(1, 32.11) = 10.09$, $p < .01$, and $F(1, 46.36) = 26.01$, $p < .001$, respectively, indicating a higher number of fixations when engaging in self-paced reading. Finally, no interaction effects were found (both cases, $Fs < 1$).

As for the number of *First-pass regressive fixations per word*, two-way mixed ANCOVA, controlling for situational interest when reading in print and prior knowledge on the texts read on the tablet, yielded a significant effect of medium, $F(1, 67) = 5.07$, $p = .03$, $\eta^2_p = .07$, showing more regressive fixations when reading on paper and a significant effect of time-frame, $F(1, 67) = 5.79$, $p = .02$, $\eta^2_p = .08$, indicating that participants who self-paced the reading time performed more regressive fixations when rereading the text. No interaction effects between medium and time-frame were found, $F < 1$. With regard to the *Second-pass regressive fixations,* two-way mixed ANCOVA, controlling for situational interest in the texts read in print, yielded no effect of reading medium, $F(1, 52) = 2.12$, $p = .15$; but a significant effect of time-frame, $F(1, 52) = 8.10$, $p < .01$, $\eta^2_p = .14$, which also revealed that participants who self-paced the reading time performed more regressive fixations on the sentences they reread. No interaction effect was found between the two factors, $F < 1$.

Finally, robust mixed ANOVA indicated that, although participants spent more time rereading the texts on paper (see Table 4.5), this effect of medium was non-significant, $F(1, 29.68) = 1.32$, $p = .26$. With regard to the effect of the reading time-frame, the participants spent significantly less time rereading the texts when reading under time pressure, $F(1, 31.92) = 19.06$, $p < .001$. Again, no interaction effect was found between the two factors, $F < 1$.

In sum, results from the eye-movement measures during first-pass reading showed that participants fixated longer and more often on the text titles when reading on paper, and they fixated more often on the text sentences on paper as well. However, there were no differences across reading media during first-pass reading with regard to fixation times on text sentences. Second-pass reading measures yielded no differences across media in any of the eye-movement measures.

4.3.3. Effects of reading medium and time-frame on metacognitive monitoring

We first conducted a two-way mixed ANCOVA with participants' *Calibration index* in each medium as within-participants factor and the time-frame as between-participants factor. The analyses yielded no effect of medium, reading time-frame, or an interaction effect, all $Fs < 1$. We also performed a two-way repeated-measures ANCOVA per medium condition, with participants' POP and scores on the comprehension questions as the two levels of the within-participants factor (i.e., calibration), and the time-frame as between-participants factor, allowing us to look at calibration more in detail within each medium. Participants' prior knowledge and interest in the text topics were included as covariates because they correlated with their POPs. Participants were overconfident, regardless of the reading time-frame, when reading on paper, $F(1, 104) = 13.72$, $p < .001$, $\eta^2_p = .12$, and when reading on the tablet, $F(1, 104) = 20.94$, $p < .001$, $\eta^2_p = 17$, but the effect size was larger when reading on the tablet. Moreover, there was no interaction effect between medium and time-frame for either medium, both Fs < 1.

Regarding participants' resolution as a group, there were noticeable differences between reading media conditions. Whereas participants' POP and their scores on the comprehension questions correlated positively when participants read in print ($r = .25$, $p < .01$, see Table 4.3), indicating that participants were somewhat accurate as a group (see Fig. 4.2, left panel), this was not the case when they read on the tablet ($r = .02$, $p = .87$, see also Fig. 4.2, right panel).

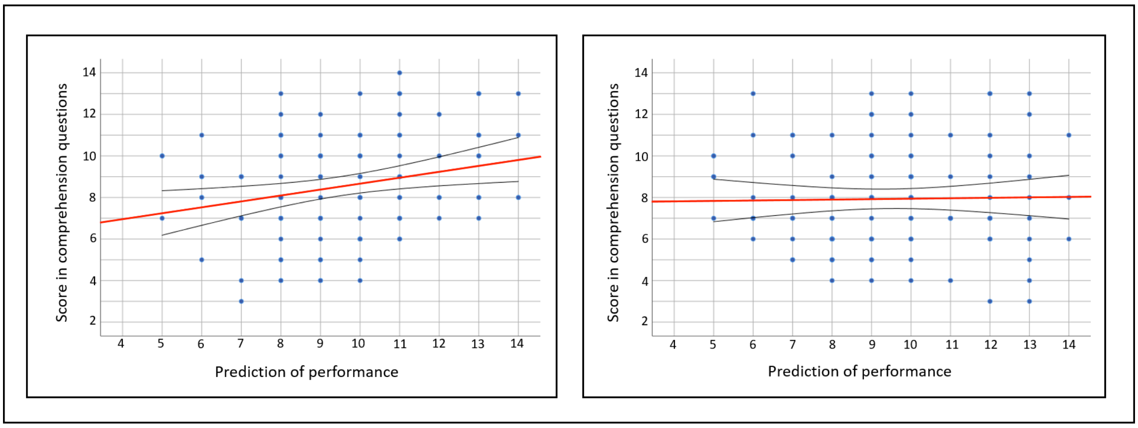

Figure 4.2. Scatter plot of the scores on the comprehension questions and the participants' POP in the on-print reading condition (left panel) and in the on-screen reading condition (right panel).

The results above, taken together, indicate that, although the size of the difference between participants' POP and actual comprehension scores was noticeably larger when reading on the tablet (η^2_p = .17 vs. η^2_p = .12), participants were overconfident to a similar extent, regardless of the reading medium and the reading time-frame (see Table 4.5). Nonetheless, there were clear differences between medium conditions in the metacognitive resolution. On this measure, participants were more accurate when reading on paper than on the tablet.

4.3.4. Effects of reading medium and time-frame on reading comprehension

Finally, we compared the scores on the comprehension questions across the experimental conditions by means of two-way mixed ANOVA. Although participants scored slightly higher in the on-print reading condition (see Table 4.5), the ANOVA results revealed that the effect of the reading medium was not significant, $F(1, 103) = 2.68, p = .10$. The effect of the reading time-frame on the comprehension scores was significant, $F(1, 103) = 7.22$, $p = < .01$, with lower scores when reading under time pressure. Finally, there was no interaction effect between the medium and the time-frame, $F < 1$.

4.4. Discussion

The present study tested the Shallowing hypothesis of on-screen reading inferiority by assessing the effect of the reading medium and reading time-frame on readers' eye movements, metacognitive monitoring, and subsequent reading comprehension outcomes. Overall, our findings only partially support this hypothesis. First, we found that the reading medium affects participants' eye movements, resulting in longer fixation times and a higher number of fixations per word on titles, and a higher number of fixations on text sentences, when reading in print than on the tablet. This pattern suggests that there is increased cognitive engagement when reading on paper, compared to reading on the tablet. Second, the relative accuracy of the participants' metacognitive monitoring was higher when reading on paper, given that their POPs and the actual performance only correlated when reading printed booklets. Lastly, although the scores on the reading comprehension questions were slightly higher when reading in print (participants scored 23.10% of an *SD* higher than when reading on the tablet), this difference did not reach significance, at least with a conservative two-tail test. Below, we shall provide more details about our results and their implications.

4.4.1. Reading medium and cognitive engagement

Regarding our first research question, the results of our study provide empirical evidence of medium effects based on participants' eye movements, showing increased engagement when reading on paper compared to reading on the tablet. The most evident differences arose when comparing text title reading across media, with longer fixation times and a higher number of fixations and regressive fixations. Increasing reading of structural parts of texts, such as headings, is considered an indicator of a more strategic reading behaviour related to overall text processing (Hyöna et al, 2002; Hyöna & Lorch, 2004; Lorch et al. 2001; Lorch & Lorch, 1996). Furthermore, readers fixated more often when reading the text sentences in the printed booklet. This effect was significant in the case of first-pass fixations, probably as a result of the higher number of first-pass regressive fixations per word. Additionally, there was a tendency toward an increased number of second-pass fixations per word. Overall, these results are in line with findings from Bansi et al. (2016), who also found increased forward fixations when reading on paper.

These results suggest that reading on the tablet hindered participants' on-task engagement. Participants were slightly more cognitively effortful when reading print, as shown by their increased fixation times and number of fixations (Rayner et al, 2006; Hyönä & Nurminen, 2006), which could also be related to a higher use of metacognitive strategies (Kinnunen & Vauras, 1995; van Gog & Jardozka, 2013). On the other hand, the fact that the difference in the number of first-pass fixations when reading the text sentences was small could explain that this difference was not found in the fixation times per word. Similar to findings from Kretzschmar et al. (2013; young adults sample), our results showed no differences in the reading times of text sentences across media.

Furthermore, we found an effect of the reading time-frame on participants' eye movements because they adapted to the time constraints by reducing their fixation time and number of fixations per word in almost every eye movement measured, except the text titles. Nevertheless, we did not replicate the interaction effect between the time-frame and the reading medium found by previous studies, which reported that the hindering effect of on-screen reading on comprehension appears especially when reading under time constraints (Ackerman & Lauterman, 2012; Delgado et al; 2018; Delgado & Salmerón, 2019). This pattern of results suggests that when difficulties appear in adapting to time limitations when reading on screens, they must be explained by other factors that cannot be captured by tracking eye movements.

4.4.2. Reading medium and metacognitive monitoring

As for our second research question, we did not find differences in the absolute accuracy of participants' metacognitive monitoring. As usually reported in previous literature, our participants were generally overconfident (see, for example, Metcalfe, 1999). Indeed, given that their POPs only correlated with their perceived prior knowledge and prior interest in text topics, they seemed to base their estimations on these two perceptions rather than on their monitoring of their level of comprehension and retention. However, unlike previous studies reporting worse metacognitive calibration when reading on screens (e.g., Ackerman & Lauterman, 2012; see also Clinton, 2019), neither the reading medium nor the time-frame influenced the overconfidence of the participants in our study.

A possible explanation for the lack of effects on calibration is that the increased overconfidence when reading on a screen found by previous investigations appears to be a consequence of a lack of ability to detect the poorer comprehension and retention in this medium. Because reading comprehension did not differ significantly in our study (as we shall discuss below), metacognitive calibration was not different across the reading media. Nevertheless, another possibility is that metacognition is not affected when reading on screens (Norman & Furnes, 2016; Singer Trakhman, Alexander, & Silverman, 2018), so that on-screen reading inferiority, when it appears, must be due to other factors related to cognitive engagement and in-depth reading. For example, the results found by

Latini et al. (2019) indicated that the participants who read an illustrated text on a computer showed less integrative processing of the textual and pictorial information than those who read the same text on paper, and they demonstrated a better integrated understanding of the text content. In another recent study (Delgado & Salmerón, 2019), on-screen readers had difficulties in reducing the occurrence of task-unrelated thoughts when the task demands (i.e., reading under time pressure) called for increasing attention. Thus, they scored lower than the on-print readers, who adapted to time constraints by reducing the frequency of their task-unrelated thoughts.

Nevertheless, although in our study absolute monitoring accuracy was not impeded when reading on the tablet, we cannot rule out the possibility that metacognitive monitoring is hampered when reading on this medium. We found differences in relative monitoring accuracy (i.e., resolution) across reading media. On the one hand, participants' POP and their actual performance on the comprehension questions correlated positively when reading the printed texts. Although the correlation index was modest, it still indicates that participants as a group were somewhat sensitive to higher and lower levels of their own comprehension and retention of the text content. On the other hand, this correlation was virtually equal to zero when the same participants read on the tablet, revealing that the ability to discriminate these levels declined notably. In short, our results suggest that metacognitive monitoring could be negatively affected when reading on the screen, although the evidence found is not strong enough to draw firm conclusions about this circumstance.

4.4.3. Reading medium and reading comprehension

The scores on the comprehension questions were slightly higher when reading on paper than when reading on the tablet. Although the difference did not reach significance with a two-tail analysis, participants in our study scored 23.10% of a *SD* higher when reading on paper than when reading on the tablet, a result that is consistent with previous findings (Clinton, 2019; Delgado et al, 2018). It could be argued that the higher level of cognitive engagement, as indicated by the differences in participants' eye movements and metacognitive resolution, resulted in higher reading comprehension in this medium. However, as mentioned above, we should not ignore the fact that this difference in comprehension between reading media was not significant, and so this conclusion should be viewed with particular caution.

In addition, there was no interaction effect between the medium and the reading time-frame on reading comprehension, so that the greater on-screen reading inferiority when reading under time constraints reported in previous studies was not replicated (Ackerman & Lauterman, 2012; Delgado et al, 2018; Delgado & Salmerón, 2019). This result was unexpected, and it is unclear why the time-frame did not influence the medium's effect in our sample. It is possible that participants in our study were not familiar with these type of tasks, so that they did not use reading strategies to adapt to reading short texts within a pressured reading time-frame. This tentative explanation could account for the lack of the effect of the medium on the participants' time management ability and other strategic decisions.

4.4.4. Limitations and further research

Our study extends previous literature on the effect of reading media on reading performance and comprehension. Findings suggest that reading on tablets hinders cognitive engagement and metacognitive monitoring, compared to reading in print. Thus, they provide some support for the Shallowing hypothesis as an explanation for the on-screen inferiority found by prior research (Clinton, 2019; Delgado et al, 2018, Kong et al., 2018). However, our results should be interpreted with caution because the differences in reading comprehension outcomes across media did not reach significance. Therefore, further research is needed to clarify the theoretical and practical implications of this issue.

One important limitation of our study is that texts might not be long enough to capture differences in readers' eye movements and their subsequent consequences for reading comprehension. Given that the greatest difference between reading media in our findings appeared when comparing the reading of text titles, further eye-tracking research should use longer and more complex texts with topic changes (e.g., Hyönä et al., 2002) and multiple headings (e.g. Bartell, Schultz, & Spyridakis, 2006). Identifying reading profiles related to the reading frequency of headings, topic-introducing sentences, and other structural parts of texts (Hyönä et al., 2002) should also be addressed by the research on reading media effects. Moreover, further investigation should study cognitive engagement by combining the eye-tracking technology with other measures such as EEG (Kretzschmar et al., 2013), mindwandering (Reichle. Reineberg, and Schooler, 2010), or postural movements (Kaakinen et al., 2018).

Additionally, although we found differences across media in the correlation scores on the comprehension questions and the participants' POP, our study did not provide the necessary number of measures to perform a more fine-grained analysis of metacognitive resolution within participants (Metcalfe & Finn, 2008). Researchers should address this issue by increasing the number of texts per medium and participant (see Ackerman & Goldsmith, 2011), or by increasing the number of predictions per text.

Finally, participants in our study read the texts, both in print and on the tablet, on reading devices that were on a stand, in order to ensure eye-tracking accuracy. Thus, the participants could not hold the devices or fully manipulate them. It has been proposed that reading is embodied (Magen & van Weel, 2016), and so differences between reading media could be caused by differences in the physical interaction between the reader and the text. Further research could explore this factor by comparing reading performance and outcomes when reading on a tablet that stands on a desk vs. a tablet held by the reader.

In short, more research is needed to further investigate not only the on-screen inferiority itself, but also the underlying mechanisms that help to explain its nature, as well as individual differences that could moderate the influence of the medium on reading performance. We believe that increasing the knowledge about this issue should be a pillar of the future of education in the Digital Era.

References

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning. *Journal of Experimental Psychology: Applied, 17*, 18-32. https://doi.org/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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154-158. https://doi.org/10.1016/j.paid.2016.02.043

Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: An eye-tracking study. *Instructional Science, 39*, 581-601. https://doi.org/10.1007/s11251-010-9142-5

Bansi, S., Oudega, M., Koornneef, A., & van den Broek, P. (2016). *The influence of presentation medium and induced beliefs on reading comprehension: An eyetracking study*. Poster session presented at the Scandinavian Workshop on Applied Eye Tracking 2016, Turku, Finland.

Baron, N. S. (2015). *Words onscreen: The fate of reading in a digital world*. New York, NY: Oxford University Press.

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 [Validation of the INFLESZ Scale to assess the readability of texts addressed to patients]. *Anales del Sistema Sanitario de Navarra, 31*, 135-152. http://scielo.isciii.es/scielo.php?pid=S1137-66272008000300004&script=sci_arttext&tlng=en

Catrysse, L., Gijbels, D., Donche, V., De Maeyer, S., Lesterhuis, M., & Van den Bossche, P. (2018). How are learning strategies reflected in the eyes? Combining results from self-reports and eye-tracking. *British Journal of Educational Psychology, 88*, 118-137. https://doi.org/10.1111/bjep.12181

Chen, D., & Catrambone, R. (2015). Paper vs. screen effects on reading comprehension, metacognition, and reader behavior. *Proceedings of The Human Factors and Ergonomics Society Annual Meeting, 59*, 332-336. https://doi.org/10.1177/1541931215591069

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, *42*, 288-325. https://doi.org/10.1111/1467-9817.12269

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenfield, S. (2015*). Mind change: How digital technologies are leaving their mark on our brains*. New York, NY: Random House Incorporated.

Delgado, P., & Salmerón, L. (2019). The inattentive on-screen reading: Reading medium affects attention and text comprehension under time pressure. Manuscript submitted for publication.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, *52*, 104-123. https://doi.org/10.1080/00461520.2017.1281747

Fiedler, K., Ackerman, R., & Scarampi, C. (2019). Metacognition: Monitoring and controlling one's own knowledge, reasoning and decisions. In R. J. Sternberg & J. Funke (Eds.), *The Psychology of Human Thought: An Introduction* (pp. 89-111). Heidelberg: Heidelberg University Publishing.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59-109. https://doi.org/10.3102/00346543074001059

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. https://doi.org/10.3758/BF03193146

Fulmer, S. M., & Frijters, J. C. (2011). Motivation during an excessively challenging reading task: The buffering role of relative topic interest. *The Journal of Experimental Education, 79*, 185–208. https://doi.org/10.1080/00220973.2010.481503

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. Bjork (Eds.), *Handbook of memory and metacognition* (pp. 411–455). Mahwah, NJ: Lawrence Erlbaum Associates.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Oxon: Routledge.

Hermena, E. W., Sheen, M., AlJassmi, M., AlFalasi, K., AlMatroushi, M., & Jordan, T. R. (2017). Reading rate and comprehension for text presented on tablet and paper: Evidence from Arabic. *Frontiers in Psychology, 8*, 257. https://doi.org/10.3389/fpsyg.2017.00257

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hynd, C. R., McWhorter, J. Y., Phares, V. L., & Suttles, C. W. (1994). The role of instructional variables in conceptual change in high school physics topics. *Journal of Research in Science Teaching*, *31*, 933-946. https://doi.org/10.1002/tea.3660310908

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 44-55. https://doi.org/10.1037/0022-0663.94.1.44

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures tostudy global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 313-334). Amsterdam: Elsevier.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354. https://doi.org/10.1037/0033-295X.87.4.329

Kaakinen, J. K., Ballenghein, U., Tissier, G., & Baccino, T. (2018). Fluctuation in cognitive engagement during reading: Evidence from concurrent recordings of postural and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 1671-1677. https://doi.org/10.1037/xlm0000539

Kaakinen, J. K., & Hyönä, J. (2005). Perspective Effects on Expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes, 40*, 239-257. https://doi.org/10.1207/s15326950dp4003_4

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse processes, 33*, 159-173. https://doi.org/10.1207/S15326950DP3302_03

Kaufman, G., & Flanagan, M. (2016). High-low split: Divergent cognitive construal levels triggered by digital and non-digital platforms. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2773-2777. https://doi.org/10.1007/s11145-003-8128-y

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, *35*, 1567-1577. https://doi.org/10.3758/BF03193491

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction, 5*, 143-165. https://doi.org/10.1016/0959-4752(95)00009-R

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363-394. https://doi.org/10.1037/0033-295X.85.5.363.

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education, 123*, 138-149. https://doi.org/10.1016/j.compedu.2018.05.005

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PLOS ONE, 8*, e56178. https://doi.org/10.1371/journal.pone.0056178

Latini, N., Bråten, I., & Salmerón, L. (2019). It's All about Sex: Reading Medium Affects Processing and Integration of Textual and Pictorial Information about Human Sexuality. Manuscript submitted for publication.

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse Processes, 54*, 427-445. https://doi.org/10.1080/0163853X.2017.1319653

Lorch, R. F., & Lorch, E. P. (1996). Effects of headings on text recall and summarization. *Contemporary Educational Psychology, 21*, 261-278. https://doi.org/10.1006/ceps.1996.0022

Lorch, R. F., Lorch, E. P. Ritchey, K. McGovern, L. Coleman, D. (2001) Effects of headings on text summarization. *Contemporary Educational Psychology 26*, 171-191. https://doi.org/10.1006/ceps.1999.1037

Mair, P., & Wilcox, R. R. (2019). *WRS2: A Collection of Robust Statistical Methods*. R package version 1.0-0. https://cran.r-project.org/package=WRS2

Mangen, A., & Van der Weel, A. (2016). The evolution of reading in the age of digitisation: an integrative framework for reading research. *Literacy*, *50*, 116-124. https://doi.org/10.1111/lit.12086

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal* *of Educational Research, 58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Metcalfe, J. (1998). Cognitive Optimism: Self-Deception or Memory-Based Processing Heuristics? *Personality and Social Psychology Review, 2*, 100-110. https://doi.org/10.1207/s15327957pspr0202_3

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*, 174-179. https://doi.org/10.3758/PBR.15.1.174

McDonald, R. (1999). *Test Theory: a Unified Treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Miller, B.W. (2015). Using reading times and eye-movements to measure cognitive engagement, *Educational Psychologist, 50*, 31-42. https://doi.org/10.1080/00461520.2015.1004068

Mizrachi, D., Salaz, A. M., Kurbanoglu, S., Boustany, J., & ARFIS Research Group. (2018). Academic reading format preferences and behaviors among university students worldwide: A comparative survey analysis. *PLOS ONE*, *13*, e0197444. https://doi.org/10.1371/journal.pone.0197444

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media? *Computers in Human Behavior, 54*, 301-309. https://doi.org/10.1016/j.chb.2015.07.043

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon, 9*(5), 1-6. https://doi.org/10.1108/10748120110424816

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory &*

Cognition, 28, 1004–1010. https://doi.org/10.3758/BF03209348

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009) Eye movements and attention in reading, scene perception, and visual search, *The Quarterly Journal of Experimental Psychology*, *62*, 1457-1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241-255. https://doi.org/10.1207/s1532799xssr1003_3

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*, 1300-1310. https://doi.org/10.1177/0956797610378686

Sackstein, S., Spark, L., & Jenkins, A. (2015). Are e-books effective tools for learning? reading speed and comprehension: iPad vs. paper. *South African Journal of Education, 35*(4), 1-14. https://doi.org/10.15700/saje.v35n4a1202

Schneider, B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-subject factors. *Frontiers in Psychology*, *6*:474. https://doi.org/10.3389/fpsyg.2015.00474

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, *13*, 23-52. https://doi.org/10.1023/A:1009004801455

Singer, L. M., & Alexander, P. A. (2017a). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*, 1007-1041. https://doi.org/10.3102/0034654317722961

Singer, L. M., & Alexander, P. A. (2017b). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education*, *85*, 155-172. https://doi.org/10.1080/00220973.2016.1143794

Singer Trakhman, L. M., Alexander, P. A., & Silverman, A. B. (2018). Profiling reading in print and digital mediums. *Learning and Instruction*, *57*, 5-17. https://doi.org/10.1016/j.learninstruc.2018.04.001

Singer Trakhman, L. M., Alexander, P. A., & Berkowitz, L. E. (2019). Effects of processing time on comprehension and calibration in print and digital mediums. *The Journal of Experimental Education, 87*, 101-115. https://doi.org/10.1080/00220973.2017.1411877

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*, 672–695. https://doi.org/10.1037/0022-3514.94.4.672

Templeton, G. F. (2011). A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research. *Communications of the Association for Information Systems*, *28.* https://doi.org/10.17705/1CAIS.02804

Thiede, K. W., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*, 66-73. https://doi.org/10.1037/0022-0663.95.1.66

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Frontiers in Psychology, 7:*769. https://doi.org/10.3389/fpsyg.2016.00769

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 832-842. https://doi.org/10.1037/a0029669

van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 143-156). New York, NY: Springer.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, *33*, 755-782. http://dx.doi.org/10.6018/analesps.33.3.268401

Walczyk, J. J. (1995). Testing a compensatory-encoding model. *Reading Research Quarterly*, 396-408. https://doi.org/10.2307/747623

Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior, 21*, 377-394. https://doi.org/10.1016/j.chb.2004.02.007

Wilcox, R. R. (2012) *Introduction to Robust Estimation and Hypothesis Testin*g (3rd ed.). Amsterdam: Elsevier.

Wolf, M. (2018). *Reader, come home: The reading brain in a digital world.* New York, NY: Harper.

Wolf, M., & Barzillai, M. (2009). The importance of deep reading. *Educational*

Leadership, 66(6), 32-37.

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.7284&rep=rep1&type=pdf

[This page intentionally left blank]

Chapter 5

Study 3

The inattentive on-screen reading: The reading medium affects attention and text comprehension under time pressure*

Pablo Delgado and Ladislao Salmerón

~

"Sometimes the mind fixes itself with so much earnestness on the contemplation of some objects, that it turns their ideas on all sides, remarks their relations and circumstances, and views every part so nicely, and with such intention, that it shuts out all other thoughts, and takes no notice of the ordinary impressions made then on the senses, which at another season would produce very sensible perceptions; at other times, it barely observes the train of ideas that succeed in the understanding, without directing and pursuing any of them; and at other times, it lets them pass almost quite unregarded, as faint shadows that make no impression."

John Locke, *An essay concerning human understanding* (1690)

* The content of this chapter is a manuscript submitted for publication.

Chapter 5

The inattentive on-screen reading: Reading medium affects attention and reading comprehension under time pressure

THIS STUDY COMPARED the influence of reading media and reading time-frame on readers' on-task attention, metacognitive calibration, and reading comprehension. One hundred and forty undergraduates were allocated to one of four experimental conditions varying in the reading medium (in print vs. on screen) and on the reading time-frame (free vs. pressured time). Readers' mindwandering while reading, prediction of performance in a comprehension test, and their text comprehension were measured. On-print readers, but not on-screen readers, mindwandered less on the pressured than in the free time condition, indicating higher task adaptation in print. Accordingly, on-screen readers in the pressured condition comprehended less than the other three groups. Mindwandering and text comprehension were similar under free reading time regardless of the medium. Lastly, there were no differences in readers' metacognitive calibration. The results support the hypothesis of shallow information processing when reading on screen under time constraints.

5.1. Introduction

Major concerns about the utility of digital technologies in education have grown as their use becomes more and more pervasive. Scholars from different disciplines are bringing up their worries about their potential harming impact on human cognition, with especial emphasis on students' in-depth information processing and sustained attention capacity (e.g, Baron, 2015; Salmerón & Delgado, 2019; Wolf, 2018). Several empirical studies have reported that the use of digital technologies at school can lead to negative learning outcomes. In light of these considerations, it seems reasonable to deem digital technologies as not always suitable for academic reading and learning.

The conclusions of three recent meta-analyses on the medium effect on reading comprehension should be a matter of concern (Clinton, 2019; Delgado, Vargas, Ackerman, & Salmerón, 2018). Results demonstrated that people comprehend less the same texts on screen than on paper. The overall effect sizes found by these studies were Hedges' $g = -.21$ (Delgado et al, 2018) and −.25 (Clinton, 2019), and analyses of moderators identified three main qualifying factors. First, both Clinton (2019) and Delgado et al. (2018) found the on-screen inferiority to be clear in expository but not narrative texts, with $g = -0.32$ (vs. $g = -0.04$) and $g = -0.27$ (vs. $g = 0.01$), respectively. Second, the effect was significant only among studies in which participants read under time constraints ($g = -.26$; Delgado et al., 2018). Finally, the effect of generation may also play a role, because the medium effect increased .01 points each year from 2001 to 2017 (i.e., the more recent the studies, the larger the on-screen inferiority; Delgado et al., 2018). Although from the classical approach of Cohen (1988) such effects are small, educational researchers have recently emphasized the need to interpret effects sizes in context (Funder & Ozer, 2019). Thus, as Delgado et al. (2018) argued, an effect size ranging from −0.21 to −0.32 is relevant in the reading comprehension field because it represents approximately 2/3 of the yearly growth in reading comprehension during elementary school (Luyten, Merrel, & Tymms, 2017).

The fact that the on-screen inferiority particularly emerges in expository texts and that it increases under time constraints suggests that such effect arises in cognitively demanding tasks. Understanding expository texts (vs. narrative) calls for a greater depth of processing, as they usually present academic knowledge by means of a large number of ideas, infrequent vocabulary and complex text structures. Conversely, narrative texts are generally simpler and closer to the real-world knowledge (Graesser & McNamara, 2011). Moreover, an increased efficiency is required when performing tasks under limited time. In such cases, in-depth processing in combination with time management becomes critical (Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). Thus, Delgado et al. (2018) pointed out to the *Shallowing hypothesis* as an explanation for on-screen inferiority (Annisette & Lafreniere, 2017). This hypothesis considers that the daily, massive experience of reading on digital media promotes a superficial way of relating with textual information, which in turn is changing the way we process information. Although this hypothesis originally refers to the way we read on any type of medium, evidence suggests that such effect is more salient when reading on screen.

Building on these empirical and theoretical backgrounds, our study seeks to disentangle the cognitive processes underlying shallow on-screen reading by analyzing undergraduate students' attention and metacognitive calibration while reading a lengthy print or digital text with or without time pressure.

5.1.1. On-task attention when reading on screen

A major concern regarding the impact of digitalization on information processing is a decreasing ability to focus on task (Baron, 2015; Wolf, 2018). From this perspective, reading on screen is inherently distracting as a result of frequent reading experiences based on skimming and multitasking. For example, Daniel and Woody (2012) found that engaging with competing activities when reading at home was more frequent among participants who read electronic versions of a textbook than among those who read it in print. Nonetheless, to the best of our knowledge, no study has directly analysed readers' attention while reading on screen, compared to reading in print. Our study is designed to fill in this gap.

On-task attention has been investigated by means of mindwandering measures. Mindwandering can be defined as unconstrained self-generated mental activity characterized by thoughts that arise independently of the task being performed, which have been called task-unrelated thoughts (TUTs; Smallwood, 2013). Mindwandering is part of a general process that implies attentional shifts from external to internal experiences. The most used method to capture the presence of TUTs is the *probe-caught technique*, where participants are periodically interrupted during the task and asked to report whether they were mindwandering. This method is considered valid and informative to assess the occurrence of TUTs (Smallwood & Schooler, 2006), and it has been used in reading research (e.g., Dixon & Bortolussi, 2013; Feng, D'Mello, & Graesser, 2013).

Reading tasks are unique to study mindwandering, because comprehending texts involves the construction of representations of the external environment (Smallwood &

Schooler, 2006). In this type of task, the occurrence and maintenance of TUTs entails that top-down attention shifts from the text content to the individual's internal activity, causing a temporary mindless reading mode. Its detrimental consequences for reading comprehension are evident (Feng et al., 2013; McVay & Kane; 2012; Soemer & Schiefele; 2019; Unsworth & McMillan, 2013).

5.1.2. Inattentive on-screen reading and the Shallowing hypothesis

The present study tests the hypothesis that shallow processing on screen is related to inattentive reading (i.e., a more frequent mindwandering). Although identifying the nature of an increasing mindwandering is beyond the scope of our study, it is helpful to explain why we expect such an effect in light of two main hypotheses for mindwandering (see Smallwood, 2013). On the one hand, the *executive failure hypothesis* assumes that mindwandering occurs due to momentary shortcomings of executive control (McVay & Kane, 2012). Therefore, if digital technologies are promoting a lack of attentional control when reading on screen, they would lead to a higher frequency of TUTs than paper, which will cause shallower processing when reading on screen.

On the other hand, the *decoupling hypothesis* states that on-task engagement and TUTs compete for the same cognitive resources (Smallwood & Schooler, 2006). Thus, the larger the share of cognitive resources dedicated to the task, the lower the mindwandering. Accordingly, shallower processing when reading on screen, whatever the cause, would itself liberate cognitive resources that could be devoted to wander. In this case, increased mindwandering would be a consequence rather than a cause of shallow reading.

In sum, the two scenarios sketched above led us to expect that the on-screen inferiority is related to higher inattentive reading. Thus, examining how reading media affect readers' mindwandering would provide a direct explanation of the on-line processes responsible of shallow processing on screen. Additionally, the effect of increased mindwandering may become more harmful when focused attention becomes more critical, such as when reading under time constraints (Ackerman & Lauterman, 2012; Delgado et al., 2018).

5.1.3. A metacognitive deficit when reading on screen and the inattentive reading

A different explanation for the shallow reading of digital texts is provided by the metacognitive deficit hypothesis. In one of the most relevant attempts to understand the underling mechanisms of on-screen inferiority, Ackerman et al. studied participants' metacognitive calibration (Ackerman & Lauterman; 2012; Lauterman & Ackerman, 2014; Sidi, Shpigelman, Zalmanov, & Ackerman, 2017). Calibration is deemed a product of self-regulated learning processes which refers to a monitoring skill that reflects the accuracy of learners' perceptions of their own performance (Pieschl, 2009). Calibration tends to be poor, with learners often being overconfident (see Stone, 2002). In a series of studies, Ackerman et al. consistently found that participants' calibration accuracy was inferior when the experimental task was accomplished on a computer relative to printed materials, both when participants read texts to answer comprehension questions (Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014), and when they solved brief problems (Sidi et al., 2017). As a consequence of this heightening metacognitive inaccuracy, the authors argued, the outcomes were poorer when performing the tasks on screen under time constraints.

Potentially, the relationship between a metacognitive monitoring deficit and inattentive reading may be bidirectional. Lower on-task attention could hinder monitoring, because during off-task periods readers cannot accurately judge, or even being aware of their current level of understanding. Conversely, overconfidence in one's level of comprehension could liberate cognitive resources which could be dedicated to mindwandering. Based on these considerations, we expect that reading on screen under time constraints will promote both higher levels of metacognitive overconfidence and mindwandering that reading on paper, regardless of the nature of the relationship between these phenomena.

5.1.4. The present study

The present study aimed to replicate the on-screen reading inferiority effect under time constraints, as well as to shed light on the explanation for such effect. Following the call for using more ecologically valid materials in reading research (Mangen, Olivier, & Velay, 2019), participants in our study read a text substantially longer than what is typical in this research field. Participants were randomly allocated to one of four experimental conditions, so that they read either in print or on screen, with or without time pressure. We measured text comprehension, mindwandering, and metacognitive calibration. Besides, we measured a comprehensive set of covariates to control for their potential influence on comprehension (Daneman & Merikle, 1996; Guthrie, Klauda, & Ho, 2013; Hidi, 2001; Ackerman & Lauterman, 2012; Naumann, 2015, Ozuru, Dempsey, & McNamara, 2009) and mindwandering (Feng et al, 2013; Fulmer, D'Mello, Strain, & Graesser, 2015; Kane & McVay, 2012; Randall, Oswald, & Beier, 2014; Unsworth & McMillan, 2013; Xu & Metcalfe; 2016). Our hypotheses were:

1. Participants reading on screen will mindwander more than those reading the

printed text, regardless of the reading time-frame.

2. Participants reading on screen under time pressure will show poorer calibration of comprehension than the other groups.
3. Participants reading on screen under time pressure will comprehend less than the other groups.

5.2. Method

5.2.1. Participants

One hundred and forty first-to-fourth year undergraduate students of pedagogy, teaching, and psychology of a large Spanish university volunteered for class credit. All participants had Spanish as their native language, and the mean age of the sample was 20.46 years (*SD* = 1.57). All participants provided informed consent, and they were debriefed after completing the study.

As indicated by a priori power analyses (G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2017) with alpha and beta levels respectively set at .05 and .20, a 140-participant sample is appropriate to detect an interaction effect of medium and time-frame both on reading comprehension and on readers' calibration respectively equal to a partial eta-squared of .07 (minimum necessary sample size = 107) and .06 (minimum necessary sample size = 125; sizes of the interactive effects found by means of a similar experimental design by Ackerman & Lauterman, 2012).

5.2.2. Materials

Text. We used a lengthy expository text[11] on human learning and artificial intelligence

[11] See appendix C

that included 3101 words and two figures, distributed across four pages. We used authentic versions of the text published in the science-dissemination magazine *Investigación y Ciencia*. In the on-print reading condition we provided the article in the actual magazine, whereas in the on-screen condition we provided its pdf version on a desktop computer (screen size 17”) (see Fig. 5.1). The pdf initially presented one page by screen, but participants were allowed to set the zoom at their own pace[12]. In case they did zoom in, scrolling down the text by using the mouse wheel was necessary.

Figure 5.1. First pages of the on-screen (pdf file) and the printed version of the reading material.

5.2.3. Dependent measures

Multiple-choice comprehension test. We constructed 21 four-alternative questions, including seven questions for each of the following three comprehension processes: text-based (i.e., a single idea explicitly stated in a single sentence), local inference (i.e., a bridging inference linking two adjacent sentences), and global inference (i.e., a bridging inference linking information located more than two sentences apart). The four response

[12] If needed, participants set the zoom in advance using the last page of the previous article.

options for each question included the target and three different distractors: near-miss (an idea located in the text that conceptually taps the target answer), thematic (a plausible answer but containing common misconceptions), and unrelated distractor (an extremely improbable answer or inconsistent with the text content) (Ozuru, Best, Bell, Witherspoon, & McNamara, 2007). We excluded two questions from each level of comprehension to increase the reliability of the scale. The omega coefficient (McDonald, 1999) was used to test the reliability, because it is considered more appropriate than Cronbach's alpha for non-ordinal items with less than 5 response options (Viladrich, Angulo-brunet, & Doval, 2017). Results indicated good reliability for the 15 questions selected (ω = .82). Thus, this test finally consisted of 15 questions (maximum score = 15).

Mindwandering probes. The frequency of TUTs was assessed by means of the probe-caught technique (Feng et al., 2013). Participants were periodically interrupted (in 99-second intervals) while reading the article to indicate whether they were paying attention to a TUT at that moment. They were previously instructed to identify on-task thoughts ("*Thoughts about the text content or about how well you are understanding it*") and TUTs ("*Thoughts about your daily stuff, a memory from the past, something in the future, your current state of being, or any other type of thought not related with the text content nor with the understanding of it*") (McKay & Kane, 2012). This measure was completed in a separated sheet of paper by ticking *yes* (*I was wandering*) or *no* (*I wasn't wandering*) for each probe. The TUT proportion on the probes was calculated for each participant, ranging from 0 to 1.

Metacognitive calibration. After reading, participants predicted their performance in the comprehension test by estimating the percentage of correct answers in a continuous 25-100% scale (Ackerman & Lauterman, 2012). Calibration for each participant was calculated by subtracting the percentage of correct answers in the comprehension test from their prediction of performance (POP), which allowed us to perform correlation analyses between this measure and the other measured variables. Besides, participants' POPs were statistically compared to the actual performance by means of repeated-measures analyses.

5.2.4. Covariates

Working memory. We used the Letter-Number Sequencing Test from the Spanish version of the Wechsler Adult Intelligence Scale-IV (Wechsler, 2008) to measure working memory capacity. In this test, the evaluator enunciates a series of alternating numbers and letters, and individuals report back the numbers from lowest to highest, and the letters in alphabetical order. Difficulty increases from a 3-item to an 8-item series. Its application procedure was adapted to a group application, and participants wrote down their responses (see Macedo-Rouet et al., 2019).

Prior knowledge. We constructed a self-reported 8-item questionnaire as an indicator of participants' prior topic knowledge[13]. Participants rated their knowledge on four subtopics related to human learning (e.g., *brain processes involved in human learning*), and four related to artificial intelligence (e.g., *computer programming*), using a scale from 1 (*I know nothing*) to 10 (*I am an expert*). Cronbach's alpha was good for the items on human learning (α = .84) and acceptable for the items on artificial intelligence (α = .70).

Topic interest. We constructed a self-reported 8-item questionnaire on participants' topic interest[14]. They rated from 1 (*not interested at all*) to 10 (*very interested*) their interest in

[13] See Appendix C
[14] See Appendix C

the same eight subtopics they rated in the prior knowledge questionnaire. The reliability was good for the items on human learning (α = .87) and acceptable for the items on artificial intelligence (α = .75).

Medium preference and use. We constructed a 5-item (two reversed) questionnaire to measure participants' medium preference for reading to learn. For each item, participants rated from 1 (*I totally disagree*) to 10 (*I totally agree*) a statement regarding the use of printed vs. digital texts for learning purposes (e.g., *I understand and memorize better when I study reading an electronic text than when I read on paper*). A mean score above 5 points indicated preference for paper. This questionnaire showed a good reliability level (α = .82). In addition, participants indicated at what age they started to use digital devices regularly, and how many hours a day they use them for leisure and for educational/professional purposes.

Perceived text difficulty and situational interest. A 2-item questionnaire required participants to rate, after reading, their perceived text difficulty from 1 (*very easy*) to 10 (*very difficult*) and their interest in the text content, also from 1 (*not interesting at all*) to 10 (*very interesting*).

5.2.5. Procedure

Tasks were completed in one small group session (six participants maximum). Sessions were conducted in a silent room and lasted approximately 75 minutes. Participants first completed the self-reported questionnaires on prior knowledge and topic interest, followed by the working memory test. Then, participants were introduced to the reading task: "*You are now going to read an article to learn as much as you can, because you will be asked to complete a test consisting of 21 four-alternative multiple-choice* *questions on the text content. Please note that you won't be allowed to go back to the text while answering the questions*". The instructions for the free-time condition continued as follows: "*You can read the article at your own pace. When you consider you have read enough, raise your hand and wait to be given the comprehension test. It is important that you do not disturb the other participants, so please do everything very silently*". The pressured-time groups had only 16 minutes and 30 seconds to read the text, which represents the 75 percent of the mean time that 20 participants in the pilot study took to read the article at their own pace. The instructions for this condition continued as follows: "*You must keep in mind that you have little time to read the article. You only have 16 minutes and a half, which is the 75 percent of the time that a group of people spent on average when reading at their own pace. I will let you know when you have gone through half the time, also when there have four minutes left, and finally when there is only one minute and a half left. You will have to stop reading when you are told that the time is up and you will then receive the questions".* Afterwards, participants were instructed in how to perform the mindwandering probe-caught task, and they were reminded to be honest when answering the probes. Participants in the pressured-time condition were probed 10 times, the last one just before reading time was over, whereas the number of probes for the participants in the free-time condition ranged from 10 to 19.

When the reading task was finished, participants predicted their performance in the reading comprehension test, and subsequently completed it. Finally, they reported their perceived text difficulty and interest in the text, and answered the questionnaire on medium preference and use.

5.3. Results

Two participants were excluded because they did not perform the task properly[15]. Data distribution of the reading comprehension scores, TUT rate, and calibration was inspected before conducting the main analyses. Eight participants were identified as outliers, because they scored in the reading comprehension test below 2 *SDs* from their group mean (sizeable differences in this measure existed between participants in the free and pressured-time groups). They were excluded from subsequent analyses, and therefore the final sample consisted of 130 participants, still appropriate according to the results of the power analyses reported above. Table 5.1 shows the distribution of participants across the experimental groups. ANOVA and chi squared analyses indicated no significant differences between the experimental groups regarding participants' age, sex, grade year, and bachelor's degree (see also Table 5.1).

Table 5.1. *Distribution of participants and participants' age, sex, grade year, and bachelor's degree across experimental condition.*

Medium	**On print**		**On screen**	
Reading time-frame	**Free**	**Pressured**	**Free**	**Pressured**
n	33	32	33	32
Age	20.39 (1.31)	20.00 (1.37)	20.82 (1.90)	20.45 (1.71)
Percentage of females	81.81	75.00	84.84	84.37
Grade year[1]:				
1st year	1	3	2	3
2nd year	12	18	11	14
3rd year	14	7	12	11
4th year	6	4	8	4
Bachelor[1]:				
Pedagogy	6	4	8	4
Psychology	5	6	2	5
School Teaching	22	22	23	23

Note. [1]In number of participants

[15] They took approximately 11 minutes to read the text, which is below three *SD* from the mean of the free time condition.

5.3.1. Covariates

Table 5.2 includes descriptive information for all measured covariates for each group, and Table 5.3 include Pearson correlations for all variables. Covariates were normally distributed after removing outliers (±2*SD*; see Kurtosis and Skewness in Table 5.2) and, thus, suitable for parametric statistical analyses, except for the medium preference measure, because most of the participants reported a strong preference for reading on paper. Regarding correlations between variables, scores on the comprehension test positively correlated with working memory and situational interest, and negatively with text difficulty. The TUT proportion positively correlated with perceived text difficulty, and negatively with participants' situational interest. The remaining covariates didn´t correlate with text comprehension scores or TUT proportion. Finally, participants' calibration score didn´t correlate with any of the covariates.

To ensure that groups were comparable, we performed a series of ANOVAs with reading medium and reading time as independent variables, and each possible covariate as dependent variable. Only two of them differed between groups. Participants under pressured time reported higher perceived text difficulty regardless of the medium, $F(1, 126) = 5.05$, $p = .03$, $\eta^2_p = .04$, and self-reported prior knowledge in human learning was higher in the on-print group than in the other groups, $F(1, 126) = 8.53$, $p < .01$, $\eta^2_p = .06$. Moreover, there was a marginal interaction effect for this variable, $F(1, 126) = 3.57$, $p = .06$, $\eta^2_p = .03$.

Table 5.2. *Means and standard deviations of the covariates in each experimental condition. Skewness and kurtosis of the covariates in the whole sample*

Medium	On print		On screen		Whole sample	
Reading time frame	Free	Pressured	Free	Pressured	Skewness[1]	Kurtosis[1]
Prior knowledge on human learning	6.57 (.94)	6.67 (.78)	6.41(.89)	5.93 (.99)	-0.17	0.23
Prior knowledge on artificial intelligence	3.04 (1.01)	3.53 (1.00)	2.92 (.83)	3.09 (.98)	0.73	-0.66
Previous interest on human learning	8.80 (.92)	8.76 (.87)	8.78 (1.10)	8.97 (.84)	0.06	-0.84
Previous interest on artificial intelligence	4.8 (1.61)	5.12 (1.26)	4.95 (1.58)	5.04 (1,59)	-0.64	-0.14
Medium preference	8.91 (1.62)	8.37 (2.01)	8.73 (2.36)	9.00 (1.39)	-2.70	9.96
Starting age of use of digital technologies	10.84 (3.39)	10.34 (2.46)	10.60 (2.96)	11.06 (2.46)	0.02	-0.32
Use of digital tech for leisure[2]	2.71 (1.78)	3.53 (2.39)	2.78 (1.21)	3.29 (1.95)	1.10	0.82
Use of digital tech for study/professional purposes[2]	4.06 (2.30)	4.65 (2.61)	4.18 (2.39)	4.89 (2.99)	-0.26	-0.72
Perceived text difficulty	6.39 (1.99)	6.94 (1.81)	6.36 (1.71)	7.11 (1.11)	-0.40	0.57
Situational interest	7.12 (2.16)	6.72 (2.33)	6.47 (2.12)	6.72 (2.24)	-0.65	-0.05
Working memory	20.97 (2.26)	21.59 (2.33)	21.67 (2.55)	21.16 (2.88)	0.44	0.17
Reading time (in min.)	24.18 (4.57)	16.30 (-)	22.49 (4.11)	16.30 (-)	0.41	0.02

Note. [1]Once outliers (±2*SD*) removed. [2]In daily hours.

Based on the differences between groups and on the correlations between covariates and dependent measures, we included working memory, situational interest, and prior knowledge in human learning as covariates in the ANCOVA for text comprehension scores. The same covariates were included in the analysis of TUT proportion. In this case, controlling for working memory was a decision driven theoretically by previous evidence (i.e., working memory and mindwandering generally correlate negatively; Kane & McVay, 2012; Unsworth & McMillan, 2013; Randall et al., 2014). Perceived text difficulty was not included as covariate in both cases due to its substantive dependence on the reading time-frame conditions (Miller & Chapman, 2001). Outlier data for each covariate were replaced by the sample mean (less than 5% of the cases for each covariate). Finally, no covariates were included in the ANOVA for metacognitive calibration, because none of the measured variables correlated with this dependent variable.

5.3.2. Effects of medium and reading time-frame on mindwandering

Means and standard deviations for the TUT proportion in each experimental condition can be seen in Table 5.4. ANCOVA assumptions with respect to normality, homogeneity of variances, and homogeneity of regression slopes between the three included covariates and the dependent variable were met. Thus, a two-way (medium x reading time) ANCOVA revealed no main effect of medium, $F(1, 123) = 1.38$, $p = .24$, and a marginal main effect of reading time-frame, $F(1, 123) = 3.50$, $p = .06$, $\eta^2_p = .03$, with a lower rate of TUTs observed for pressured-time groups. A marginally significant interactive effect of media and reading time-frame was also revealed, $F(1, 123) = 4.26$, $p = .06$, $\eta^2_p = .03$. Post hoc Bonferroni-corrected pairwise comparisons indicated that, across time-frames, participants reported a significant lower TUT proportion in the on-print group under time pressure than under free-time, $F(1, 123) = 7.06$, $p < .01$, $\eta^2_p = .05$, whereas it was not statistically different across reading time groups in the on-screen condition, $F<1$. Moreover, across reading media, the TUT proportion was significantly lower in the on-print group than in the on-screen group when reading under time pressure, $F(1, 123) = 4.36$, $p = .04$, $\eta^2_p = .03$. There was no difference between media under free reading time, $F<1$.

Table 5.3. *Pearson correlations between all the measured variables.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Prior knowledge on artificial intelligence	-													
2. Prior knowledge on human learning	.25**	-												
3. Previous interest on artificial intelligence	-48**	-.01	-											
4. Previous interest on human learning	-.12	.10	.01	-										
5. Medium preference[1]	-.09	.03	.00	.11	-									
6. Starting age of use of digital technologies	.08	.08	.09	.03	-.01	-								
7. Daily use of digital tech for leisure	.06	-.20*	.02	-.19*	-.17	-.14	-							
8. Daily use of digital tech for study/prof.	.15	-.14	-.05	-.07	-.18*	-.06	.24**	-						
9. Perceived text difficulty	-.04	-.06	-.18	-.06	.24**	.08	-.04	.00	-					
10. Situational interest	.05	.02	.03	.01	-.04	.11	-.15	.04	-.10	-				
11. Working memory	.20*	.03	.24**	-.12	-.11	.10	.06	.11	-.12	.23**	-			
12. TUT proportion	-.04	-.01	-.14	-.04	-.03	-.05	.04	-.05	.21*	-.40**	-.03	-		
13. Calibration	.07	.03	-.08	.05	-.04	.16	-.02	.05	-.12	-.09	-.16	.06	-	
14. Comprehension questions	.01	.16	-.01	-.12	-.10	-.13	.04	-.08	-.19*	.28**	.21*	-.24**	-.68**	-
	1	2	3	4	5[1]	6	7	8	9	10	11	12	13	14

Note. * $p < .05$; ** $p < .01$; [1] Spearman's rho correlations.

Tablet 5.4. *Means (SD) for TUT proportion, POPs, and reading comprehension scores. Skewness and Kurtosis for each dependent variable.*

Medium	On print		On screen		Whole Sample	
Reading time-frame	Free	Pressured	Free	Pressured	Skewness[1]	Kurtosis[1]
TUT proportion	.22 (.15)	.15 (.14)	.22 (.15)	.21 (.14)	0.45	-0.59
POPs in comprehension test[2]	64.55 (12.28)	66.41 (12.91)	62.76 (11.50)	55.06 (15.88)	-0.24	-0.67
Text comprehension[3]	64.84 (15.75)	62.71 (19.00)	66.87(13.38)	53.13(17.64)	-0.26	-0.36

Note. [1]Once outliers (±2*SD*) removed. [2]Estimated percentage of correct answers. [3]Percentage of correct answers.

5.3.3. Effects of medium and reading time-frame on calibration

Participants' POPs for the comprehension test were compared to their actual performance to examine whether their metacognitive calibration differed across experimental groups. We conducted a two-way repeated measures mixed ANOVA with medium and time-frame as between-participants factors, and calibration as a within-participants factor (i.e., participants' POPs vs. text comprehension scores). Given that calibration is our focus here, we report only results from the tests of within-participants effects. Results revealed no differences between POPs and actual performance in the whole sample, $F<1$, and no interaction effects between calibration and medium, $F<1$, calibration and reading time-frame, $F(1, 126) = 2.69$, $p = .10$, and calibration, medium and time-frame, $F<1$ (see also Table 5.4). Thus, participants showed to be well calibrated regardless of the reading medium and the reading time-frame.

5.3.4. Effects of medium and reading time-frame on reading comprehension

Finally, differences in text comprehension scores were examined by means of a two-way

ANCOVA. Means and standard deviations for each group can be seen in Table 5.4. ANCOVA assumptions with respect to normality, homogeneity of variances, and homogeneity of regression slopes between the included covariates and the dependent variable were met in all cases except for participants' situational interest. Thus, we applied the blocking procedure (Tabachnick & Fidell, 2014) to include this covariate as an independent variable, because it allowed to focus on the effects of the independent variables of interest (i.e., medium and time-frame) once the variation of this covariate is removed from the estimated error. Following Tabarchnick and Fidell (2014) procedure, a new independent variable was created by categorizing situational interest values into three levels (low, medium, high) based on percentiles of 33 and 66. We then performed a 3-way ANCOVA, with medium, time-frame, and situational interest as independent variables, and working memory and prior knowledge on human learning as covariates.

The results revealed a non-significant effect of medium, $F<1$, but a significant main effect of reading time-frame, $F(1, 116) = 5.85$, $p = .02$, $\eta^2_p = .05$, indicating higher scores in the free time condition. A marginally significant interaction qualified these effects, $F(1, 116) = 3.65$, $p = .06$, $\eta^2_p = .03$. Across time-frames, post-hoc Bonferroni-corrected pairwise comparisons showed that in the on-screen condition, participants who read under time pressure scored significantly lower than those who read at their own pace, $F(1, 116) = 8.74$, $p < .01$, $\eta^2_p = .07$. This was not the case for the on-print condition, where participants scored similarly regardless of the reading time-frame, $F<1$. Furthermore, across reading media, participants who read under time pressure on screen scored significantly lower than those who read in print, $F(1, 116) = 4.17$, $p = .04$, $\eta^2_p = .04$, whereas no medium difference was found for participants reading with free time, $F<1$.

5.4. Discussion

The present investigation examined for the first time how reading medium affects readers' on-task attention while reading an authentic, lengthy expository text. It also contributed to the research efforts on how medium affects reading comprehension and metacognitive calibration. Controlling for a comprehensive set of covariates, our findings revealed that reading on screen prevented readers from reducing their mindwandering when the task requirements called for efficient reading. Accordingly, participants reading on screen under time pressure scored significantly lower in the reading comprehension test, relative to those in the other three groups. Finally, contrary to our expectations, participants were equally well calibrated regardless of the experimental group. Next, we discuss the implications of these results with respect to the influence of mindwandering on text comprehension, and how the lack of increased attention when reading on-screen can explain the screen-inferiority effect.

5.4.1. On-task attention and on-screen reading comprehension

Our experimental design provided a direct test of two potential underlying factors of the screen inferiority effect: inattentive reading and metacognitive calibration deficit. With respect to our first hypothesis, we expected to observe inattentive reading (i.e., higher TUT rate) on screen, as compared to on-print reading, regardless of the time-frame. As a consequence of this increased mindwandering, we hypothesized that disruption in reading comprehension due to mindwandering would be especially noticeable under time pressure. Yet, inattentive on-screen reading, although as such confirmed by our results, emerged in a different manner. When participants read at their own pace, they mindwandered to a similar extent, regardless of the medium. But when they read with time constraints, only on-print participants reduced the frequency of TUTs. Previous evidence has shown that learners can control the occurrence of mindwandering when the task demands call for it, especially those with greater working memory capacity (Rummel & Boywitt, 2014; Smallwood & Andrews-Hanna, 2013). Thus, given that in the present study participants' working memory did not differ across groups, we should have observed also reduced mindwandering when reading on screen under time pressure. Nevertheless, our results indicate that on-screen readers struggle to adjust to such high task demands. Although the size of this effect in our study was small according to Cohen's benchmark (η^2_p = .03, equivalent to *d* = 0.35), it is larger than the effect of an increased mindwandering when reading difficult (vs. easy) texts found by previous studies –*OR* = 1.24, equivalent to *d* = 0.12 (Feng et al., 2013); R^2 = .016 equivalent to *d* = 0.25 (Mills, D'Mello, & Kopp, 2015)–. Similarly, Latini, Bråten, Anmarkrud, and Salmerón (2019) recently reported that on-print readers were more able than on-screen readers to adapt to the learning demands of a multiple document comprehension task. Specifically, when instructed to prepare for an exam, as opposed to reading for pleasure, on-print readers wrote longer essays and indirectly integrated better the information from different sources. This adaptive strategy was not present in the on-screen readers.

In line with previous findings (Feng et al., 2013; McVay & Kane; 2012; Soemer & Schiefele; 2019; Unsworth & McMillan, 2013), the TUT proportion in our study correlated negatively with text comprehension scores. Results from the on-print group indicated that participants reduced their mindwandering while reading under time pressure, as compared to when reading with free time. This accommodation of readers' attention could have counteracted the detrimental effect of time pressure, resulting in similar comprehension scores across the two on-print groups. However, that was not the case for participants who read on screen. They did not adaptively decrease the frequency of TUTs, and their reading comprehension scores were significantly lower. This result is in line with previous findings showing on-screen inferiority when reading under time constraints (Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014).

In sum, our findings partially support the inattentive reading hypothesis by pointing to difficulties observed by on-screen readers to meet the task demands calling for increased on-task attention. Accordingly, the fact that on-screen participants in the pressured-time group didn´t reduce their mindwandering provided direct evidence for a shallow processing during on-screen reading (Annisette & Lafreniere, 2017).

5.4.2. On-screen reading and metacognitive monitoring

Unexpectedly, our results showed that undergraduates could accurately calibrate their reading comprehension regardless of the reading medium and the time-frame, contrary to previous findings yielding better calibration when reading in print (e.g., Ackerman & Lauterman, 2012). It was also unexpected that our participants were well calibrated in all the experimental conditions, because it has been widely reported that learners tend to be overconfident (Stone, 2002). The experimental procedure employed in the present study may have helped participants to make accurate predictions about their level of performance. Firstly, the available reading time in the pressured-time condition was certainly scarce, as indicated by the fact that approximately half of the participants in these groups could not reach the end of the text. Secondly, the caught-probe technique could make participants aware of their distraction while reading. Both factors may have been used by participants as cues to identify to what extent they could complete the reading assignment. The observed significant negative correlation between the TUT proportion and POPs supports this idea. Therefore, these two circumstances could have lead participants to be cautious rather than overconfident in their POPs. Thus, in spite of our results, the metacognitive deficit observed in prior studies for on-screen reading can´t be discarded. So far, whether on-screen reading harms metacognitive calibration, and under what circumstances it occurs, if so, still remains an open question (Singer Trakhman, Alexander, & Silverman, 2018).

5.4.3. Inattentive on-screen reading and theories of mindwandering

Although the aim of this study was not to identify the nature of increased mindwandering during on screen reading, our results also provided important insights on this issue. The fact that participants in the on-screen reading condition could not reduce their mindwandering when the task called for it, as the on-print readers did, indicated that screens prompt a cognitive engagement characterized by diminished attentional control. Thus, it could be argued that our findings support the executive failure hypothesis (McVay & Kane, 2012) as an explanation for the occurrence of TUTs. Nonetheless, investigating why and how mindwandering increases when reading on screen is in the core of understanding the mindwandering phenomenon itself, so we call researches to further investigate this issue in order to improve our knowledge not only about how we read on screens, but also about how and why our mind wanders especially during on-screen reading.

Additionally, the deficit in metacognitive monitoring when performing tasks on screen, especially under time pressure, reported in previous research (Ackerman & Lauterman, 2012; Clinton, 2019; Lauterman & Ackerman, 2014; Sidi et al., 2017), could also be related to difficulties in reducing mindwandering. There is no reason to consider that this deficit is constrained to metacognitive judgments. On the contrary, mindwandering could affect other monitoring processes related to task engaging, such as meta-consciousness. Several studies have examined the metacognitive status of mindwandering tracking participants' awareness of their TUTs. According to the meta-awareness hypothesis (Schooler, 2002), results showed that readers are often unaware of their mindwandering (Smallwood, 2013; Smallwood & Schooler, 2006). Thus, a broader metacognitive deficit could lead not only to an increased calibration inaccuracy, but also to generate TUTs more often when reading on screen than when reading printed texts. This possibility could be tested in future studies.

5.4.4. Limitations and future research

The present study is not exempt from limitations. Although using an authentic, lengthy text is a strength of our study, we can´t generalize our results to shorter texts. Further research should test if the inattentive reading hypothesis can explain screen inferiority in shorter learning tasks (cf. Sidi et al, 2017). In addition, although our study included a comprehensive set of covariate measures, other individual factors, such as participants' sustained attention capacity, or a more exhaustive working memory measurement, could help to further explain mindwandering during on-screen reading. Future studies could investigate whether the on-screen inferiority effect depends on those individual differences, to the extent that they can be related to the inattentive reading as an explanation (cf. Ben-Yehudah, & Brann, 2019). Furthermore, this research line could be especially relevant in school-aged students, because they show a larger variability with respect to individual differences.

It is also noteworthy that we can´t rule out the possibility that reading on screen is not inattentive in itself, but its disruptive effect may be caused by the fact that reading on desktop screens involves a body position that could hinder on-task attention. Reading is deemed a cognitive activity with an embodied nature, and the physical relationship between reader and text is different between these media (Mangen & van der Weel, 2016). In this vein, the regular body posture used while reading printed materials on a desk could facilitate the immersion in the text, increasing on-task attention. To rule out this option, our study could be replicated by using tablets instead of desktop computers, given that hand-held devices allow for a reader-text physical interaction more similar to that of reading on paper.

5.5. Conclusion and educational implications

Our results showed that reading on screen lead to inattentive reading particularly when the task demands an increase in on-task attention for efficient information processing. We argue that this inattentive reading is causing, at least in part, a shallow information processing and lower comprehension. Our findings support current concerns indicating that digital technologies, under certain circumstances, hinder reading and learning (Baron, 2015; Salmerón & Delgado, 2019; Wolf, 2018). As argued above, although the size of the on-screen inferiority effect under time pressure found in our study was small according to Cohen's (1988) benchmarks (η^2_p = .04, analyzing it in context provides a richer picture. This effect size is located at the lower bound of Hattie's 'zone of desired effects' in educational contexts (i.e., a medium effect; Hattie, 2009), and it represents slightly more than the yearly growth in reading comprehension during elementary school (0.32; Luyten et al., 2017). Therefore, this result should be a matter of concern to educational practitioners and policy makers. There are clear educational scenarios where on-screen reading, in light of our results, should be avoided. For example, taking exams on screen could prevent students from fully demonstrating their knowledge and skills, because they may struggle to adjust their attentional focus to their full potential. But suggesting a ban on digital technologies for educational purposes would be naïve in the XXI century. They are here to stay and they offer a wide range of educational possibilities. Nevertheless, we call educational practitioners and policymakers to consider the fact that printed texts are more appropriate when it comes to in-depth reading, especially with lengthy texts. In this regard, educational systems should be especially cautious with recent campaigns supporting a complete shift from printed to e-text books. Instead, we should find an appropriate balance between the use of printed materials and digital technologies by means of evidence-based decisions. This is a major goal for education in the Digital Age.

References

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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154-158. https://doi.org/10.1016/j.paid.2016.02.043

Baron, N. S. (2015). *Words onscreen: The fate of reading in a digital world*. New York, NY: Oxford University Press.

Ben-Yehudah, G., & Brann, A. (2019). Pay attention to digital text: The impact of the media on text comprehension and self-monitoring in higher-education students with ADHD. *Research in Developmental Disabilities, 89,* 120-129. https://doi.org/10.1016/j.ridd.2019.04.001

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, *42*, 288-325. https://doi.org/10.1111/1467-9817.12269

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433. https://doi.org/10.3758/BF03214546

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. https://doi.org/10.1016/j.edurev.2018.09.003

Dixon, P., & Bortolussi, M. (2013). Construction, integration, and mind wandering in reading. *Canadian Journal of Experimental Psychology, 67*, 1-10. https://doi.org/10.1037/a0031234

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. https://doi.org/10.3758/BF03193146

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review, 20*, 586-592. https://doi.org/10.3758/s13423-012-0367-y

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology*, *41*, 98-110. https://doi.org/10.1016/j.cedpsych.2014.12.005

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*, 156-168. https://doi.org/10.1177/2515245919847202

Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modelling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly, 48*, 9-26. https://doi.org/10.1002/rrq.035

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to* *achievement*. Oxon: Routledge.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*, 191-209. https://doi.org/10.1023/A:101666762

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*, 348-354. https://doi.org/10.1177/0963721412454875

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high-and low-achieving primary school children's reading. *Learning and Instruction*, *5*, 143-165. https://doi.org/10.1016/0959-4752(95)00009-R

Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019) Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology, 59.* https://doi.org/10.1016/j.cedpsych.2019.101797

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Luyten, H., Merrell, C., & Tymms, P. (2017). The contribution of schooling to learning gains of pupils in Years 1 to 6. *School Effectiveness and School Improvement*, *28*, 374-405. https://doi.org/10.1080/09243453.2017.1297312.

Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadtler, M., Salmerón, L., & Rouet, J-F. (2019). How good is this page? Benefits and limits of prompting on adolescents' evaluation of Web information quality. *Reading Research Quarterly,* *54*, 299-321. https://doi.org/10.1002/rrq.241

Mangen, A., Olivier, G., & Velay, J. L. (2019). Comparing comprehension of a long text read in print book and on Kindle: Where in the text and when in the story? *Frontiers in Psychology, 10:* 38. https://doi.org/10.3389/fpsyg.2019.00038

Mangen, A., & van der Weel, A. (2016). The evolution of reading in the age of digitisation: an integrative framework for reading research. *Literacy*, *50*, 116-124. https://doi.org/10.1111/lit.12086

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141*, 302-320. https://doi.org/10.1037/a0025250

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*, 40-48. https://doi.org/10.1037/0021-843X.110.1.40

Mills, C., D'Mello, S. K., & Kopp, K. (2015). The influence of consequence value and text difficulty on affect, attention, and learning while reading instructional texts. *Learning and Instruction*, *40*, 9-20. https://doi.org/10.1016/j.learninstruc.2015.07.003

Naumann, J. (2015). A model of online reading engagement: linking engagement, navigation, and performance in digital reading. *Computers in Human Behavior, 53,* 263-277. https://doi.org/10.1016/j.chb.2015.06.051

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.

https://doi.org/10.1080/07370000701632371

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*, 228-242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Pieschl, S. (2009). Metacognitive calibration—an extended conceptualization and potential applications. *Metacognition and Learning, 4*, 3-31. https://doi.org/10.1007/s11409-008-9030-4

Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mindwandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin, 140*, 1411-1431. https://doi.org/10.1037/a0037428

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mindwandering to situational demands. *Psychonomic Bulletin & Review*, *21*, 1309-1315. https://doi.org/10.3758/s13423-013-0580-3

Salmerón, L., & Delgado, P. (2019). Critical analysis of the effects of the digital technologies on reading and learning/Análisis crítico sobre los efectos de las tecnologías digitales en la lectura y el aprendizaje. *Cultura y Educación*, *31*, 465-480. https://doi.org/10.1080/11356405.2019.1630958

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences, 6*, 339-344. https://doi.org/10.1016/S1364-6613(02)01949-6

Sidi, Y., Shpigelman, M., Zalmanov, H. y Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction, 51*, 61-73.

https://doi.org/10.1016/j.learninstruc.2017.01.002

Singer Trakhman, L. M., Alexander, P. A., & Silverman, A. B. (2018). Profiling reading in print and digital mediums. *Learning and Instruction*, *57*, 5-17. https://doi.org/10.1016/j.learninstruc.2018.04.001

Smallwood, J. (2013). Distinguishing how from why the mind wanders: a process–occurrence framework for self-generated mental activity. *Psychological Bulletin, 139*, 519-535. https://doi.org/10.1037/a0030010

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mindwandering state. *Frontiers in Psychology*, *4*, 441. https://doi.org/10.3389/fpsyg.2013.00441

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*, 946-958. https://doi.org/0.1037/0033-2909.132.6.946

Soemer, A., & Schiefele U. (2019). Text difficulty, topic interest, and mind wandering during reading. *Learning and Instruction, 61,* 12-21. https://doi.org/10.1016/j.learninstruc.2018.12.006

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*, 437-475. https://doi.org/10.1023/A:1009084430926

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics*. Harlow: Pearson Education.

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 832-842. https://doi.org/10.1037/a0029669

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, *33*, 755-782. http://dx.doi.org/10.6018/analesps.33.3.268401

Wechsler, D. (2008) *Wechsler Adult Intelligence Scale – IV.* San Antonio, TX: Pearson.

Wolf, M. (2018). *Reader, come home: The reading brain in a digital world.* New York, NY: Harper.

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*, 681-695. https://doi.org/10.3758/s13421-016-0589-8

[This page intentionally left blank]

Chapter 6

General Discussion

~

"Knowledge that can predict is better than knowledge that merely records, because it helps us better in the main business for which knowledge exists–to control the forces of nature and ourselves."

Edward L. Thorndike, *Education: A first book* (1912)

Chapter 6

General discussion

NEEDLESS TO SAY, the burgeoning presence of digital technologies in educational practices has been unstoppable. Nevertheless, as we have argued over the present dissertation, the learning outcomes of the use of these technologies as educational tools are still controversial. In chapters 1 and 2, we reviewed some findings revealing that the use of digital technologies, especially computers, seems to be not always appropriate for learning in educational settings. Within that broad landscape, we have shown in chapters 3 to 5 that, specifically, the use of these digital devices as reading tools seems to be detrimental for reading performance. Recent warnings about this circumstance are indeed not scarce (e.g., Baron, 2015; Carr, 2010; Mangen & van der Weel, 2016; Wolf, 2018). In this final chapter, we shall first summarize the main conclusions drawn from this dissertation. Then, we shall reflect on future directions in research on the reading medium effect and we shall discuss on the main educational implications of our findings.

The main goal of the present dissertation was twofold. First, we aimed to synthesize the existing literature about the reading medium effect on reading outcomes, and to clarify this effect. Second, we tried to expand empirical evidence on the reading medium effect which also helped testing the Shallowing hypothesis as an explanation for it. To that ends, we first conducted a meta-analysis that included primary studies comparing reading comprehension between reading on paper and reading on screen (Study 1, see Chapter 3). Despite the heterogeneity of the effect sizes included in the analyses, we found an overall mean effect size favouring on-print reading both among the between-participants studies and among the within-participants studies, which goes in line with findings by other recent meta-analyses published almost at the same time (Clinton, 2019; Kong et al., 2018). Additionally, we based on several substantive, methodological, and extrinsic variables to qualify the included studies. We found that the on-screen inferiority was larger among those studies that posed a limited reading time to participants (vs. self-paced reading time) and among those that used expository texts (vs. narrative texts), which also goes in line with the finding by Clinton (2019). Finally, we found that the more recent the studies, the larger the medium effect.

Two main conclusions were drawn from our meta-analysis. On the one hand, the main, obvious conclusion is that reading on screen prevents readers from making the most of their reading comprehension ability. We noted that the size of the effect, although could deemed as small according to general classifications such as Cohen's (1988), approximately represents two thirds of the mean yearly growth in reading comprehension achievement during the elementary school (Luyten, Merrell, & Tymms, 2017). In other words, it means that the level of reading comprehension of an average elementary student tested during the last weeks of an academic year could decrease to the level that he or she initially had at the begging of the academic year if tested on screen. Indeed, as we mentioned in Study 3 (see Chapter 5), the size of this effect is located at the lower bound of the *teacher effects zone* of Hattie's (2009) classification, which represents the range of effect sizes that formal instruction typically achieves in an academic year. On the other hand, findings regarding the three moderators that qualified the on-screen inferiority effect are in line with the explanation for this effect proposed by the Shallowing hypothesis. As we have argued, the increased effect both when reading under limited time, and when reading expository texts points out that the detrimental influence of on-screen reading becomes more salient when the task demands for increased cognitive efforts. Moreover, the fact that the medium effect increases among the more recent studies suggests that the on-screen inferiority is larger in younger generations. This finding is also in line with the Shallowing hypothesis, because this hypothesis proposes that the medium effect is a consequence of the habitual relaxed way in which people consume information within digital environments (typically the Internet). Accordingly, the more the people access information in this shallow manner when reading on screen, the larger the medium effect on reading performance. Hence, on-screen inferiority will be larger in those who have grown surrounded by a more massive daily digital experience, namely, the younger generations.

Building on the above results, we carried out two empirical studies that aimed to experimentally test the Shallowing hypothesis by comparing readers' eye movements (Study 2, see Chapter 4) and readers' mindwandering while reading (Study 3, see Chapter 5), respectively. Additionally, both studies compared readers' metacognitive monitoring of the text comprehension across reading media, as well as they examined the influence of the reading time-frame on the medium effect. Overall, findings from both studies support the hypothesis that reading on screen hinders readers' cognitive engagement. In Study 2 we found that readers fixated longer and more often on texts' titles, and also fixated more often on texts' sentences when reading on print. Additionally, readers' metacognitive resolution at group level was more accurate when reading in this medium. Accordingly, readers' comprehension scores were slightly higher in the on-print condition, although the difference was not significant with a conservative two-tail analysis (p = .10). Finally, we found no influence of the reading time-frame on the medium effect, but this factor was decisive in Study 3, as we shall summarize next.

Findings from Study 3 showed that only those participants who read on print (vs. on screen) reduced their mindwandering when the reading time was constrained. In other words, on-screen readers did not increase on-task cognitive engagement when the situation called for it, as the on-print readers did. Thus, on the one hand, reading comprehension when reading on screen was affected by reading time-frame, as on-screen participants who read under time pressure scored lower than those who self-paced their reading time and, more important, than those who read under time pressure in print. On the other hand, on-print readers scored similarly regardless of the time frame. Thus, the medium effect on readers' cognitive engagement arose in those participants who read under time pressure, which replicates previous findings (see Study 1; see also Ackerman & Lauterman, 2012). However, in this case there were no differences in participants' metacognitive monitoring across reading media. Participants in all the experimental conditions were indeed quite well accurate, which differs from the overestimation consistently reported in previous literature (see Stone, 2002). As we have argued, a possible explanation for this rare high monitoring accuracy could be due that participants in our study could have based their predictions of performance on their level of mindwandering, given that the probe-caught technique required them to self-record the occurrence of task-unrelated thoughts (TUTs) at each probe.

Why the reading time-frame did not moderate the reading medium effect in Study 2 remains unclear, but, as we argued, it is possible that the experimental materials and procedure did not allow participants' decisions about cognitive effort allocation to make a difference. As a tentative explanation, we suggested that our participants may not be familiar with reading tasks in which short texts are read within a pressuring reading time, so that they did not adapted their strategic reading to this situation. Thus, decisions on strategic processing influencing the reading process did not take place (such as, for example, reduction of mindwandering, selective skimming, or rereading decisions).

In short, the studies presented in this dissertation yield two main conclusions that make this research work particularly valuable. Firstly, the results from the meta-analysis in combination with those from our two experimental studies consistently point out an on-screen reading inferiority when it comes to comprehending texts –see also Clinton (2019) and Kong et al. (2018)–, which seems to specially arise when reading under time constraints. Notably, this medium effect has also been recently found by large-scale studies evaluating not only reading comprehension (Jerrim, Micklewright, Heine, Salzer, & McKeown, 2018), but also mathematics and science achievement (Fishbein, Martin, Mullis, & Foy, 2019). These latter findings suggest that the hindering influence of the digital medium is a cognitive effect that goes beyond reading and extends to other academic subjects and tasks (see also Sidi et al., 2017). Secondly, Study 2 is one of the few experimental studies comparing the investment of cognitive effort across reading media by measuring eye movements (see also Bansi et al, 2016; Kretzschmar et al., 2013; Latini, Bråten, & Salmerón, 2019) and, to our knowledge, Study 3 is the first experimental cross-media comparison of on-task sustained attention. Our results are consistent with the prediction of a cognitive disengagement when reading on screen compared to reading on paper, so they support the Shallowing hypothesis as an explanation for the medium effect.

Our conclusions are in line with previous findings that also indicate a lower level of cognitive and metacognitive engagement when reading on screen (Ackerman & Lauterman, 2012; Kaufman & Flanagan, 2016; Latini, Bråten, & Salmerón, 2019; Lauterman & Ackerman, 2014). Kaufman and Flanagan (2016) reported that high-level comprehension and the use of high-level strategies were higher in print when reading and solving problems, respectively. Furthermore, Ackerman et al. found that participants' metacognitive monitoring under time pressure was more accurate in print also when reading (Ackerman & Lauterman, 2012) and when solving problems (Sidi et al., 2017). Finally, the recent eye-tracking study conducted by Latini, Bråten, and Salmerón (2019) found that on-print readers showed more gaze transitions between textual and pictorial information when reading an illustrated text. These authors concluded that reading on screen hindered readers' cognitive processes of integration of both informational formats. Thus, except for the eye-movement and EEG results from Kretzschmar et al. (2013), who concluded similar cognitive efforts across media, our results and those mentioned above altogether represent first empirical evidences in support of the Shallowing hypothesis.

6.1. Reflections for further research

Although our results shed some light on the reading medium effect, there is still much more to research on this issue. Given the heterogeneity of the results reported both in our meta-analysis and in previous research syntheses (Dillon, 1992; Kingston, 2008; Singer & Alexander, 2017; Wang et al, 2007), it is still necessary to further clarify the circumstances that foster on-screen inferiority. Reading is a complex and multidimensional cognitive activity that consequently requires a comprehensive research approach. In this sense, we agree with Mangen and van der Weel (2016) when they call for a comprehensive account of digital reading, grounded in an integrative framework including of different research methods and complementary disciplines into our empirical tradition. We next discuss about possible future research lines on the effect of medium on reading performance. Based on previous models of reading comprehension (Snow, 2002; Rouet et al., 2017), we shall distinguish between variables related to the reader (i.e., individual differences), the text, the task demands, and the context (see Table 6.1) that may be related to the medium effect. With regard to reader-related factors our findings seem to indicate that some individual skills related to cognitive engagement are affected by the medium (i.e., attention, metacognition, and maybe the use of reading strategies), which, in turn, affects reading comprehension. Thus, on the one hand, these factors can be considered as mediators of the reading medium effect[16]. On the other hand, it is reasonable to suggest that these individual skills, at the same time, would make readers more or less liable to be affected by the medium effect.

Moreover, according to the Shallowing hypothesis, we consider factors such as the type and frequency of use of the digital technologies (both at individual and at contextual level) as the determinants of the reading medium effect. The remaining variables are suggested to be moderators of this effect (see Fig 6.1). For analytic purposes, we shall discuss about future research on these types of variables separately, but it should be borne in mind that it is their mutual interaction that will likely interact with the medium

[16] We also propose readers' standards of coherence as a possible mediator (see below)

effect. Finally, we shall briefly discuss about the integration of different research methods and about some possible interventions aiming to overcome the on-screen inferiority.

Table 6.1. *Relevant variables for further research on the reading medium effect.*

Dimensions	Variables
The reader	Frequency and type of use of digital technologies Attentional capacity Metacognitive skills Use of reading strategies Standards of coherence
The text	Structural elements (headings, subheadings, snippets) Pictures or graphical information The presence of interesting but irrelevant information (e.g., seductive details, advertisements) Need for scrolling in digital text
The task	Reading time-frame Task goals/Reading purposes
The context	Macro-level context: national, home, and school amount and type of uses of digital techs. Micro-level context: • Presence of competing activities. • Digital context: Text on the Internet vs. text on screen

6.1.1. The reader and the medium effect.

As proposed by the Shallowing hypothesis, the type and frequency of use of the digital technologies as reading devices are individual factors that determine the medium effect. Furthermore, we have argued that metacognition (Study 2; see also Ackerman & Lauterman, 2012; Clinton, 2019), sustained attention (Study 3), and maybe the use of reading strategies (Study 2; see also Latini, Bråten, & Salmerón, 2019) are reader-related skills that seems to be negatively affected when reading on screen, so leading to poorer reading comprehension. Hence, the degree of individual development of these skills could, in turn, moderate the medium effect. We next discuss further research regarding these set of variables. For the sake of simplicity, the role of readers' standards of coherence shall be discussed in a following section, in relation to task goals and purposes.

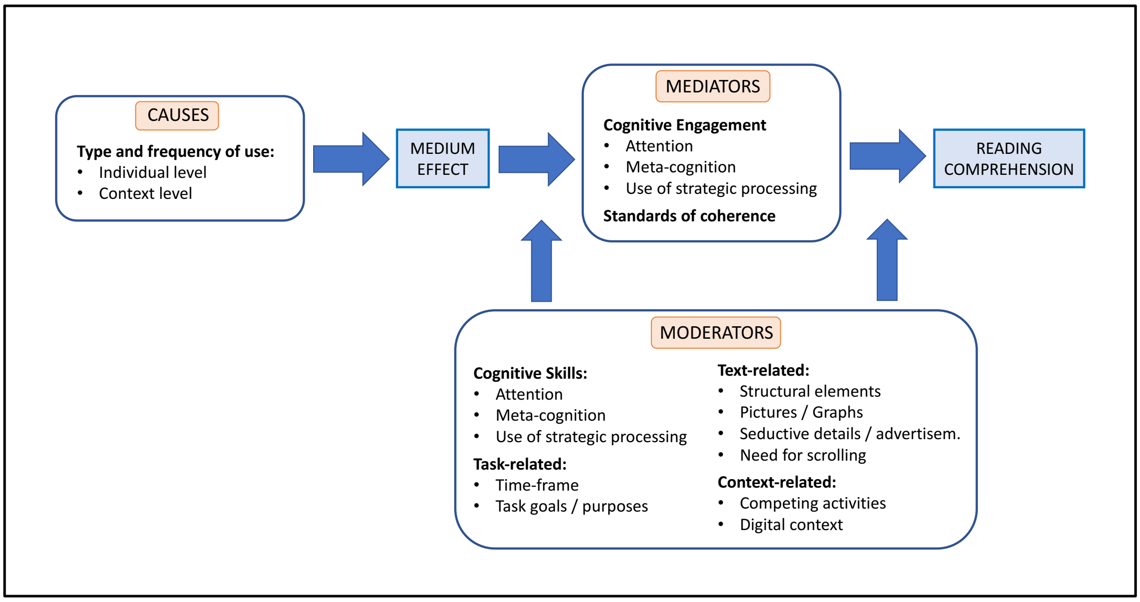

Figure 6.1. Proposed influences of the variables related to the medium effect.

The frequency and type of use of digital technologies are clearly interrelated individual variables that likely have much to say. In this direction, some studies have found that social media use and texting habits correlate negatively with student's engagement in reflective thought (Annisette & Lafreniere, 2017) and with self-reported distraction during academic reading (Levine et al., 2007). According to the Shallowing hypothesis, a habitual reading of short textual pieces of information by means of quick interactions on digital environments would lead to a reading habit characterized by a lack of in-depth processing. A recent study by Macedo-Rouet et al. (2019) found that the frequency of use of social network sites by teenagers is negatively related to their skills to select reliable sources when reading conflicting texts on the same topic. Thus, it would be expected that the more the digital devices are used for reading with leisure purposes or for reading through webpage navigation, the larger the medium effect. Future longitudinal studies on this issue could explore how long-term use of digital devices for leisure reading and specially the use social media sites could be related to the reading medium effect and its developmental course.

Moreover, as mentioned in Chapter 1, digital technologies are associated with an increased daily multi-tasking activity (e.g., Rideout, 2015, see Greenfield, 2009) which, in turn, seems to be related to deficits in cognitive skills (e.g., Ophir et al., 2009; Sanbonmatsu et al., 2013) and academic achievement (e.g., Carrier et al., 2015; van der Schuur et. al, 2015). We hence suggest that the influence of multitasking habits on the medium effect should also be explored. One possibility is to categorize participants varying in the amount of daily media multitasking to examine the medium effect across categories. This could also be explored by means of longitudinal designs in which individuals' multitasking habits serve as possible predictors of the medium effect. Another possibility could be to examine the short-term effect of multitasking by priming a group of participants a kind of *multitasking mind-setting* by asking them to perform some multitasking activity prior to the reading task. The medium effect on this group could be compared to a non-prompted group –see Kaufman & Flanagan (2016, Study 3B) for a similar design–. It would be expected that prompting a multitasking mind-setting leads on-print readers to perform similar to on-screen readers, because it would foster a superficial processing regardless of the medium.

With regard to the moderating effect of other reader-related factors, we have argued that reading on-screen affects readers' metacognitive monitoring and the allocation of on-task attentional effort, so it is plausible to consider that the degree of development of these cognitive capacities play a relevant role in the medium effect. Therefore, it could be expected that readers low in sustained attention capacity or in metacognitive skills will be harmed to a greater extent by reading on screen. Further research could test this possibility by comparing the reading medium effect across different groups of readers varying in these individual skills. For example, Ben-Yehudah and Brann (2019) have recently found that adults with ADHD were negatively affected by reading on screen to a greater extent than their normative-developed peers.

Finally, the use of different reading strategies may also play an important role. As reported in Study 2, we found that readers fixated longer and more often on texts titles when reading on the printed booklet. In this direction, as mentioned above, Latini, Bråten, and Salmerón (2019) found an increased strategic integrative processing when reading in print, as compared to reading on screen. Therefore, if the reading medium affects the use of reading strategies, it should be examined whether different profiles of readers interact with the reading medium. As mentioned in the Study 2, Hyönä et al. (2002) measured readers' eye movements and they found that those who invested more time to read the headings and final sentences of each topic[17] constructed better summaries of the texts content. They called this profile of readers as *topic structure processors.* Moreover, Hyöna et al. (2002) also found that this type of readers had the largest working memory

[17] Participants read two expository texts, each including 12 topics.

capacity. In this regard, the results from our Study 2 seem to indicate that reading on screen prevents readers from using this type of strategies related to a deeper text comprehension. Future investigations could adapt the Hyönä and colleague's study to compare the use of different processing strategies across reading media. It could also be explored whether different cognitive capacities, such as working memory or sustained attention, moderate the use of reading strategies across reading media. Given that the Shallowing hypothesis considers that on-screen reading hinders cognitive engagement, we would expect not only a lower number of topic processors on screen than on paper, but also a lack of relation (or a lower one) between individual differences and strategic processing on screen, as compared to reading on paper. Another possibility is that those higher in the use of in-depth processing strategies will be less affected by the medium effect.

6.1.2. Texts characteristics and the medium effect

Text characteristics are assumed to exert a large influence on reading comprehension (e.g., Kintsch, 1988, 1998). Among the possible texts features that could be related to the on-screen reading inferiority, we shall discuss here about how the presence of structural elements, (such as headings), pictorial information, distracting elements, and the need for scrolling could qualify this effect.

As argued above, it is possible that the on-screen reading hinders in-depth processing of structural elements, but the interaction between these structural elements and the reading medium can also be investigated in a reversed way, namely, aiming to examine whether the presence of these elements affects reading performance differently across media. Bartell, Schultz, and Spyridakis (2006) reported that the frequency of headings in texts influenced text comprehension only in those participants who read on screen. Their results showed that a low and a high frequency of headings were detrimental for on-screen readers' comprehension, as compared to a medium frequency, whereas on-print readers' comprehension was similar regardless of this text feature. We still have to gain more knowledge about possible differences across media in the processing of different structural elements, different types of text structures (e.g., Bohn-Gettler & Kendeou, 2014), and other types of textual elements (e.g., the presence of snippets presenting relevant information from the text). Methods combining off-line measures and online variables, such as readers' eye movements, provide unique data to reach these goals. We expect that strategic processing of the text elements or structurally relevant parts would be higher when reading printed texts.

Other text elements than structural ones also deserve future research effort to examine how reading on screen affects reading performance. For example, whether pictures or graphics adding complementary information are processed differently across media remains almost unexplored. To the extent of our knowledge, only the eye-tracking study by Latini, Bråten, and Salmerón (2019) have examined this possibility by using an illustrated text. Besides the need for extending their findings on differences in processing graphical elements, we suggest to explore readers' on-task cognitive engagement by introducing the presence of *seductive details*, such as interesting but irrelevant pictures (Mayer, 2003), or the presence of static advertisements[18]. Given that findings from Study 3 indicated that on-screen readers did not increased on-task attention when the situation

[18] We specify *static* advertisements due to the existence of other type of formats that are specific to digital environments, such as pop-ups windows or dynamic advertisements, whose influence cannot be explored in the printed format.

called for it, it could be explored whether these types of irrelevant elements are more or less distractive while reading depending on the reading medium. The influence of the reading time-frame could also be investigated. Based on the Shallowing hypothesis and given our results about the effect of the medium on sustained attention, we suggest that readers would avoid distracting elements to a greater extent when reading on paper, which would be especially salient when reading under time pressure.

Finally, how texts are delivered on screen could also affect reading comprehension. As we found in the meta-analysis (Study 1), the on-screen inferiority effect tended to be larger when the texts on screen had to be scrolled than when the whole pages were presented one by one. Only Pommerich (2004) and Higgins et al. (2005) experimentally tested the influence of this digital feature, and they also found that comprehension tended to be poorer when texts had to be scrolled. However, the difference did not reach significance in both cases. Given these results, if the need for scrolling harms readers' comprehension by obstructing their processing of the visual structure of the text, a larger detrimental effect of this digital feature would be expected when reading long texts. Thus, this possibility should be explored by using longer expository texts (e.g., textbook chapters, scientific papers) than those usually used in reading research (typically less than 1000 words length), as was the case of the study by Higgins et al. (2005)[19].

6.1.3. Task demands and the medium effect

As we have showed, reading under time pressure seems to increase the on-screen inferiority. Still, we consider necessary to continue extending research on the influence

[19] Unfortunately, no information about the texts length was reported in Pommerich (2004). Neither in Higgins et al. (2005), but it was provided by the author following a personal request.

of the reading time-frame. Designing complex reading tasks, in which time management becomes crucial to accomplish the task goals, would expand evidence on this effect. For example, instead of asking participants to read single texts, which is the usual procedure in research on the medium effect, the interaction between the reading time-frame and the medium could be further explored in a multiple-document context, in which gathering and integrating information from different texts is necessary to meet the task demands (e.g., Latini, Bråten, Anmarkrud, et al., 2019). Additionally, text relevance could also be manipulated across documents in order to increase task complexity by requiring participants to make decisions about what texts they should use, or to what extent they have to be read in depth. We suggest that tasks requiring higher integrative processing and selection of relevant information will increase the medium effect, especially when performing the task under time constraints. It would be expected that readers show better and more selective integrative processing of different documents when reading in print, and that differences across media will be lower when the task requires the comprehension of information conveyed in single texts.

The medium effect and the Shallowing hypothesis could be also tested in relation to task goals by rooting the research in the Landscape model of reading comprehension, specifically, into the *standards of coherence* framework. As mentioned in Chapter 1, standards of coherence are considered the reader's desired level of comprehension in a given reading task, and they are assumed to guide attention-allocation and the use of different reading strategies (Butterfuss & Kendeou, 2018; van den Broek et al., 2011; van den Broek & Helder, 2017). Although the standards of coherence represent an individual factor, they are strongly dependent of task goals and/or reading purposes. Thus, if reading on screen fosters a shallower comprehension of the text, we could expect that this effect affects readers' standards of coherence. We propose two possible ways in which it could occur. On the one hand, the digital medium could prompt readers to assume a lower standard of coherence than the printed medium, given the assumption that readers usually read quickly and superficially on screen. Therefore, readers would devote less cognitive efforts when reading on screen. On the other hand, the digital medium could lead readers to wrongly estimate they have attained the desired standard of coherence. Notably, the latter possibility is in line with the increased readers' overconfidence in their level of comprehension when reading on screen (see Clinton, 2019). However, we see these two possibilities as complementary rather than competing hypothesis, whose predictions can be tested by means of an experimental design that manipulates readers' standards of coherence by prompting different reading goals.

In this line, the study by Latini, Bråten, Anmarkrud, et al. (2019) is the first attempt to explore cross-media differences in the influence of reading purposes (for pleasure vs. for an exam) on participants' comprehension and integration of the content of two texts on the same topic. After reading the texts, participants had to answer four comprehension questions. Results showed that reading purpose indirectly influenced text integration through the length of the responses to the comprehension questions only when reading the printed texts. Thus, on-print participants who read to study showed better text integration than those who read for pleasure. This effect was not found in the on-screen participants, who did not adapt to reading purposes. In terms of standards of coherence, it could be argued that readers adapted standards to purposes only when reading in print. However, there were no differences in text comprehension and integration across reading media. The study by Latini, Bråten, Anmarkrud, et al. (2019) represents a first step into what we consider a quite profitable field of research. Future investigations could extend their findings by replicating this study also requiring participants to read within a constrained time-frame. Moreover, other ways to manipulate reading goals could also be explored. For example, readers could be asked to imagine that they have to learn about some issue in order to solve take a decision relevant to their own lives, or to accomplish some working goals instead of passing an exam (e.g., McCrudden, Magliano, & Schraw, 2010). Furthermore, the interaction between medium and reading purposes should be examined in an ecological, actual classroom environment, in which different tasks vary in whether they ask for reading, for pleasure, or for learning about the subject at hand.

6.1.4. The context and the reading medium

The reading contexts can be studied at macro and at micro level. On the one hand, the contextual macro level refers to sociocultural practices that occur in social groups or communities (Snow, 2002). In this regard, the influence of macro-level variables, such as the penetration of digital technologies at a national level, the prevalence of different types of use at homes and of reading habits, or the types of instructional practices at school could be explored by large-scale studies comparing reading performance across media. To what extent this type of large-scale macro-level indicators influence the reading medium effect is still unknown. Analysing data from international studies such as PISA or PIRLS[20] could provide valuable insights to increase our knowledge on these factors.

On the other hand, the context could also be considered at a micro level. According to the RESOLV model (Rouet et al., 2017), the reading context also refers to

[20] The Progress in International Reading Literacy Study (PIRLS), is conducted by the International Association for the Evaluation of Educational Achievement, and provide cross-national data on children reading achievement.

the immediate contextual clues that influences readers' decisions through their "initial representation of contextual demands and opportunities" (p. 200). As we synthesised in Chapter 1, this initial representation is what authors call a *context schema*, which it is assumed to be turned into "a set of initial goals and actions" (Rouet et al., 2017, p. 200) that are also included into the *task model* (i.e., the reader's interpretation of the goals and actions demanded by the task). For instance, facing a reading task in the classroom potentially activates a context schema that triggers initial goals and decision different to those activated when reading on the Internet within a leisure context.

A contextual characteristic that may moderate the medium effect is the amount of competing activities that are present while reading. As mentioned in previous chapters, Daniel and Woody (2012) found that competing activities at home, such as accessing social media or interacting with relatives or home mates, are more distractive when reading on screen. Thus, it is possible that the medium effect is moderated by the usual level of presence of distractors within each context. Reading on digital devices would be even more detrimental at home than at the classroom, as the former usually presents more distractors than the latter. Further, the influence of distractors could also be explored by comparing between different reading purposes, as well as its interactions with individual factors, such as working memory or sustained attention capacity (e.g., comparing ADHD individuals with normative-developed peers).

Additionally, the Internet can be considered a virtual context that could activate a context schema that is different to the schema that is activated when readers merely access a text on screen, for example, when the texts are delivered in a computer by opening a pdf file. According to the Shallowing hypothesis, the way in which information is delivered and accessed on the Internet, (characterized by short texts, a random access to information through hyperlinks, and a huge amount of distracting elements such as pictures and multiple links to other type of information) is promoting the shallow way in which we read on this medium. Thus, it could be the case that, within a learning context, reading on screen a pdf file or a text that is part of an e-textbook activates a context schema that triggers a deeper processing than reading the same text by accessing a webpage. Future research could explore this possibility, also in combination with task goals and individual differences.

6.1.5. Methodologies to explore the reading medium effect

As argued above, reading is a complex and multidimensional activity in which different cognitive processes take place in interaction with external factors (i.e., the task, the text, and the context). Examining the effect of the reading medium on off-line measures as well as on isolated on-line measures is a valuable starting point that provides important findings. Nonetheless, the results discussed over the present dissertation still yield an incomplete picture. A wide range of questions related to the nature of the medium effect still remain unanswered. For example, we do not know whether the medium effect on mindwandering is a consequence of a specific inattentive mind-setting or it is caused by an overall diminished executive function activity. In this case, combining mindwandering measures with other on-line measures such as EEG could provide an reliable answer. Questions like this requires research on multiple on-line measures that synchronizes data from eye movements, EEG, mind-wandering, the use of reading strategies and metacognitive processes, and other indicators of cognitive engagement such as postural movement (e.g., Kaakinen, Ballengheim, Tissier, & Baccino, 2018) or even physiological responses. For example, with regard to this latter type of measures, it has been reported that heart rate decreases during periods of sustained attention (e.g., Lansink & Richards, 1997; Porges, & Raskin; 1969). Interestingly, a recent study by Enea, Maresh, Miller, Pritzl, and Trieglaff (2019) found that heart rate increased in the participants who read a text on their smartphones, but decreased in those who read the same text in print. Although no differences were found in a two-week delayed test on memory for text details, their heart rates seem to indicate that reading on smartphones hindered on-task sustained attention. Thus, combining these type of measures with others such as mindwandering, eye-movements, or EEG data would provide a more complete picture.

Given the cognitive disengagement when reading on screen proposed by the Shallowing hypothesis, and that the findings from Study 3 indicated difficulties in increasing on-task attention when reading on this medium, the use of techniques measuring mindwandering still have much to say in research on reading media. For example, the on-screen metacognitive deficit previously found (see Clinton, 2019) could also be related to an increased generation and maintenance of TUTs when reading on this medium (see Conclusions in Study 3). Furthermore, it could be examined whether the reading medium affects the relationship between mindwandering and task difficulty. The occurrence of TUTs decreases with moderately difficult tasks, as performers invest more cognitive efforts to task solving –although it increases with more difficult tasks, likely as a consequence of disengagement and frustration (e.g., Randall, Beier, & Villado, 2019; Xu & Metcalfe, 2016)–. Given the results from our Study 3, only participants in the on-print condition showed that trend, as they reduced their mindwandering when task difficulty increased due to reading time constraints. Thus, the fact that on-screen readers did not decrease their mindwandering when task difficulty moderately increased deserves further examinations. Is it a mere matter of attentional difficulties (as we argued in Study

3), or is it a consequence of readers being less sensitive to the increased task difficulty when reading on screen? Given previous findings indicating a metacognitive deficit when reading on screen, we suggest that these two possible explanations are also complementary rather than mutually exclusive.

Moreover, mind-wandering could vary across different digital devices, such as desktop/laptop computers, tablets, smartphones and e-reader. On the one hand, as argued in Study 3, it is possible that other digital devices than computers (i.e., hand-held devices) facilitate on-task attention, because they may foster closer embodied human-device interaction (Mangen & van der Weel, 2016). If so, reading on hand-held devices may favour on-task sustained attention, as compared to reading on computers. On the other hand, tablets and smartphones are massively used for accessing social media, multitasking, and different types of apps and games, so it is also reasonable to expect that they also activate an inattentive mind-setting. Further experimental studies should test these two predictions. In this sense, we suggest that e-readers, which are devices mostly used for reading single long texts, could offer a reading experience cognitively more similar to paper.

As mentioned above, eye tracking measures could be used to explore how readers deal with structural elements across media (Hyönä et al., 2002; Latini et al., 2019), as well as to examine differences in the effect of distractors such as seductive details or even advertisements (e.g., Hervet, Guérard, Tremblay, & Chtourou, 2010). In this regard, it is important to increase research combining eye-tracking measures with other online measures such as EEG (e.g., Kretzschmar et al., 2013) or postural changes –i.e., reader's body movements that reflect the level of cognitive engagement (e.g., Kaakinen et al., 2018)–, because it yields a richer picture of reading performance. Synchronizing different on-line measures would allow to undercover whether the medium effect is limited to specific cognitive skills (e.g., attention, metacognitive monitoring) or is affecting reading processes in a broader sense.

Additionally, tasks that measure the level of depth of information processing could be used to study the medium effect. For example, with regard to the processing and integration of a coherent representation of the text into the reader's mental model, Hannon and Daneman (2004) conducted an interesting study in the field of individual differences in reading comprehension, which could be adapted to test the Shallowing hypothesis. They developed a set of four short passages based on Barton and Stanford's (1993) *anomaly detection task* to investigate the deep of semantic processing across skilled and less-skilled readers. The original task presents the reader one of four versions of a short passage which ends with one of four local semantic inconsistences that are relatively difficult to detect (Barton & Stanford, 1993), so that it demands a scrupulous processing of the information at textbase level (Kitsch, 1988, 1998). Based on this task, Hannon and Daneman (2004, Experiment 2) created four new passages each containing on of the four types of anomalies, which allowed a within-participants design. Their findings replicated those from Barton and Stanford (1993), as they showed that readers often failed to detect anomalies, which led the authors to conclude that readers usually "assume coherence as a default as long as there is sufficient global coherence" (p. 202). Moreover, Hannon and Daneman (2004, Experiment 2) found that less-skilled readers were significantly more susceptible to failures in detecting the anomalies, so they concluded that reading comprehension skill is negatively associated with a shallow semantic processing. Based on their work, the Shallowing hypothesis predicts that reading on screen would led readers to fail detecting anomalies to a greater extent than reading in print.

As mentioned above, we agree with Mangen and van der Weel (2016) on the need for an integrative and interdisciplinary framework accounting for the digital reading reality. In this sense, conclusions drawn from phenomenology research on people's habits when interacting with digital technologies could be highly valuable to test the Shallowing hypothesis (e.g., Aagaard et al., 2015; Rose, 2011; Rowsell, 2014). Reading has changed not only because of the medium itself, but also due to what readers can do with their hands and to the existence of a new wide range of different ways of interacting with texts. Nowadays texts can be clicked, swiped, tapped, or zoomed. How these new behaviours relate to the reading act as an embodied cognitive activity has been studied by the phenomenology of reading, but it represents a research perspective mostly ignored by cognitive research. These are single- or few-case studies that yield a detailed description that undoubtedly would enrich conclusions from our empirical, experimental works. An example of what phenomenology could provide to our research realm is provided by Rowsell (2014). She observed a child called Jeremy while using an iPad to read aloud a monomodal text, to read aloud an interactive text, and to play a game (*Minecraft*). This is how Rowsell describes what she observed:

> "What I noted as Jeremy worked through the monomodal eBook was a reader struggling to get through a chapter book. Jeremy was not struggling because he could not decode words or understand meanings, rather he seemed distracted and had trouble focusing. In particular, Jeremy read in a monotone voice and evinced little affect during reading. Then, I observed Jeremy read a Marvel comic, *Iron Man*, in an interactive iBook format. There was some elevation in terms of interest and affect, but it was marginal compared with the notable change in his affect when he played *Minecraft*" (p. 123).

The example above is just a short piece that illustrates a child who is not comfortable when using an iPad to read a traditional text, and whose interest and engagement increase as the content delivered by the iPad is closer to what digital devices usually offer. To what extent these interactions and behavioural changes are related to different degrees of comprehension is still unknown. We strongly support the idea that our experimental tradition should take advantage of these detailed descriptions of the embodied reading act. Combining these observational methodologies with empirical data from on-line reading processes and off-line outcomes will help know more in depth the causes of the reading medium effect and the circumstances that foster it.

Finally, for the sake of a high-quality research activity, beyond the decisive role played by the study designs, we would like to highlight the importance of being as informative and transparent as possible when reporting investigations. There are growing concerns about the need for increasing transparency in social, cognitive, and behavioural sciences (e.g., Miguel et al., 2014; Munafò et al., 2017; Nosek et al., 2015; Pashler & Wagenmakers, 2012). Scholars widely assume that openness and replicability are fundamental values, however, these features are not commonly met by research practices (John, Loewenstein, & Prelec, 2012). In this direction, it is noteworthy that most of the studies included in our meta-analysis lacked of relevant information regarding their methods and even basic statistical data such as group mean scores and standard deviations (see Table A1 and Table A2 in Appendix A). We found missing information about at least one of the variables analysed as possible moderators in 79% of the reports of the studies. Although fortunately some authors provided the requested information, still there was a considerable amount of cases in which we lacked information that we considered important to qualify the on-screen inferiority. Thus, most of the analyses of moderating effects did not include every study from our dataset. This situation, also warned by Singer and Alexander (2017), is especially worrying when it comes to studies published in peer-reviewed journals, as the generalizability and applicability of research outcomes in educational practices is at stake.

6.2. Educational implications and interventions to overcome the medium effect

6.2.1. The use of digital devices for reading in education

As we have emphasized in previous chapters, the main educational implication of our findings is that reading printed texts seems to be more suitable than reading on screen when aiming an in-depth comprehension and learning from texts. Within those reading situations in which increased cognitive efforts are required, the digital medium seems to prevent readers from fully engaging with the reading task. As a consequence, educational practitioners could put students at risk of comprehending less if the texts are delivered on digital devices. Specifically, our results indicate that the on-screen inferiority especially appears when reading under circumstances that students usually face within educational settings, such as reading and learning from expository texts and reading under time constraints. Thus, educators should be sensible to these findings when making instructional decisions.

However, we do not call for banning digital technologies from education, as it would be not only absurd, but especially an ominous idea. For example, as mentioned in Chapter 1 with regard to readers' impressions, digital environments provide easy and quick access to different documents (e.g., Liew et al., 2000) and electronic texts are sometimes preferred because they are searchable and cost-saving (e.g., Ji et al., 2014;

Mizrachi et al., 2018). Moreover, we agree with Coscarelli and Coiro (2014) when they say that "we do not need to blame computers for destroying people and society" and that "computers, and more specifically the Internet, provides new digital spaces within which people can access diverse texts while practicing ways of becoming critical readers, autonomous apprentices, and better informed global citizens" (pp. 756-757). We also assume that digital reading goes beyond replicating texts on screen and that this offers expanding learning possibilities. For example, as pointed out by Dobler (2015) with regard to the adoption of e-textbooks in classrooms, these tools have changed the nature of texts, as they offer a new range of educational possibilities. She describes *enhanced e-textbooks* as those that give readers additional options by presenting information through various media (e.g., video, audio, hyperlinks to information on the Internet). Nevertheless, we should also be cautious about the presence and use of these digital affordances. For example, a meta-analysis on the use of electronic storybooks by young children revealed that complementary animated pictures enhance children's text comprehension. However, other affordances which imply multitasking, such as hypermedia interactive features, were found to be detrimental (Takacs, Swart & Bus, 2015). The use of digital affordances while reading represents additional decision-making situations that may harm reading comprehension. For instance, it has been argued that reading hypertexts by navigating through hyperlinks may be more cognitively challenging and may lead readers to lose the track of the textual content, which could be especially harmful for readers with low working memory and low prior knowledge (see DeStefano & LeFevre, 2007).

With regard to the use of videos for learning, although the results from research on this issue are still inconclusive, some studies have shown that videos may be less suitable for learning than texts in some cases. Recently, Lee and List (2019) and Salmerón, Sampietro, and Delgado (2020) asked a sample of undergraduates and of elementary students, respectively, to read or watch two documents delivered either in written form or on video. In both cases, results indicated that the integration of information across the two sources was better when the documents were presented in text format than in video format. List and Ballenger (2019) conducted a similar study in which they presented two documents varying in the type of content (expository and narrative). In this case, they found better comprehension of the expository document as a text than as a video, whereas comprehension of the narrative document did no differ across formats. Thus, although other studies have found no differences in comprehension regardless of the presentation format (List, 2018) and even better comprehension for single documents when they were presented on video (Lee & List, 2019), results seem to indicate that learning from videos may be detrimental when it comes to high demanding cognitive processes, such as integrative comprehension.

Still, we have no doubt that the additional features and advantages provided by digital environments and electronic texts may be fruitful for learning as long as its use is based on well-founded instructional designs. For example, according to Mayer (2014), presenting information by means of different modalities favors learning if they complement and enrich each other, and they avoid cognitive overload. Nevertheless, as we have argued over the present dissertation, they seem to be not always appropriate for learning. Almost two decades ago scholars started to talk about the *digital native* as a new type of learner. In Prensky's (2001) seminal article, the author stated that the digital natives prefer to learn by means of what digital environments offer them. As he argued, they are good at accessing information randomly and at multi-tasking. Based on these ideas, some scholars have called for instructional practices that meet the digital native's preferences and inherent digital skills (e.g., Palfrey & Gasser, 2008). However, a growing literature corpus claims that the digital native learners do not exist. It seems that younger generations are not especially good at learning through those ways that were supposed to be preferred by and profitable for them, such us multi-tasking and quick and random interactions with multiple pieces of information (see Kirchsner & De Bruyckere, 2017; Margaryan, Littlejohn, & Vojt, 2011), which could be even detrimental (e.g., Sanbonmatsu et al., 2013; Ophir et al., 2009; van der Schuur et al., 2015). In fact, Prensky himself later suggested to "think in terms of digital wisdom" rather than distinguishing between digital natives and digital immigrants, and he called for a "wisdom in the prudent use of technology to enhance our capabilities" (2009, p. 1).

Within the specific field of reading, we have showed evidence also contradicting the existence of the digital native as a learner skilled at reading and learning on screen. We have argued that younger generations are indeed more sensible to the hindering effect of on-screen reading on text comprehension. We should not ignore what research on the use of digital technologies in education, particularly, as reading tools is revealing. Thus, we adhere to the Stavanger Declaration (E-READ, 2019), which reports the main conclusions and recommendations drawn by a European collaborative research project on reading in digital contexts[21]. This declaration "concerning the future of reading" states that

> "Paper and screens each afford their own types of processing. In today's hybrid reading environment of paper and screens, we will need to find the best ways

[21] E-READ: The evolution of reading in the age of digitisation. Results from our meta-analysis (Chapter 3) represent some of the key findings of this project.

> to utilize the advantages of both paper and digital technologies across age groups and purposes. [..] Teachers and other educators must be made aware that rapid and indiscriminate swaps of print, paper, and pencils for digital technologies in primary education are not neutral. Unless accompanied by carefully developed digital learning tools and strategies, they may cause a setback in the development of children's reading comprehension and emerging critical thinking skills" (E-READ, 2019).

6.2.2. Helping students to overcome the medium effect.

The Stavanger Declaration also recommends that "students should be taught strategies they can use to master deep reading and higher-level reading processes on digital devices". Thus, tasks aiming to encourage in-depth processing of the information could be valuable intervention tools to overcome the on-screen inferiority. However, this is also an unexplored avenue. The study by Lauterman and Ackerman (2014) is an exception in this realm. They tested the effect of practicing a delayed keywords-generation task that has been previously found to foster meta-comprehension (e.g., Thiede, Dunlosky, Griffin, & Wiley, 2005). After reading each text, participants were asked to write down four keywords capturing the gist of the text. Interestingly, this task improved on-screen readers' metacognitive monitoring and reading comprehension, so that their reading performance was similar to on-print readers. Nevertheless, this procedure was more effective for those on-screen readers who declared that they preferred studying on screen (Lauterman & Ackerman, 2014).

Another way to promote in-depth processing that could help overcome the on-screen inferiority is the generation of self-explanations. The term self-explaining "refers to the activity of generating explanations to oneself, usually in the context of learning from an expository text" Chi, 2000, pp. 164-165). This activity is considered "a process that the learner uses to help himself or herself understand external inputs that can be instantiated by any medium", whose "focus is on trying to understand the learning materials and make sense of it" (p. 169). Importantly, practicing self-explanations has also demonstrated to foster monitoring one's own level of comprehension (e.g., Griffin, Wiley, & Thiede, 2008), as well as to help generate inferences and to improve the reader's mental model of the text, especially when reading expository texts (see Chi, 2000). Thus, given that reading expository texts has been found to be more sensible to the reading medium effect (Study 1; see also Clinton, 2019), we suggest future research to test whether this learning activity is useful to help readers avoid the shallower on-screen reading.

However, although we propose that activities inducing in-depth processing could help overcome the on-screen inferiority, these tasks should be carefully selected and designed. For example, text-annotating, an activity that apparently could also increase cognitive engagement on screen, has been found to be fruitful only when reading in print. Accordingly, differences across reading media increases when participants were asked to make annotations (Ben-Yehudah & Eshet-Alkalai, 2014). Still, although the latter finding indicates that more research is needed to support evidence-founded decisions in this field, it is possible that some ways of prompting in-depth processing are beneficial for increasing comprehension and learning when reading on screen.

6.3. A few final words

In short, the present dissertation has yielded empirical evidence indicating that reading on screen may lead to poorer comprehension, especially when reading long expository texts under time constraints. Moreover, we have provided some support for the Shallowing hypothesis as an explanation for this medium effect. Our results from readers' eye movements, metacognitive resolution, and mindwandering led us to conclude that reading on screen prevents readers from fully engaging with the reading task. Education cannot wait and see about digital technologies anymore. Finding a balanced use of digital technologies and training students to fully benefit from them should be two crucial goals for our educational systems. Thus, we honestly hope that our work has provided relevant results and conclusions that help support an evidence-based educational use of digital technologies.

References

Aagaard, J. (2015). Drawn to distraction: A qualitative study of off-task use of educational technology. *Computers & Education*, *87*, 90-97. https://doi.org/10.1016/j.compedu.2015.03.010

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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Annisette, L.E., & Lafreniere, K.D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154-158. https://doi.org/10.1016/j.paid.2016.02.043

Bansi, S., Oudega, M., Koornneef, A., & van den Broek, P. (2016). *The influence of presentation medium and induced beliefs on reading comprehension: An eyetracking study*. Poster session presented at the Scandinavian Workshop on Applied Eye Tracking 2016, Turku, Finland.

Baron, N. S. (2015). *Words onscreen: The fate of reading in a digital world*. Oxford: Oxford University Press.

Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition, 21,* 477-487. https://doi.org/10.3758/BF03197179

Bartell, A. L., Schultz, L. D., & Spyridakis, J. H. (2006). The effect of heading frequency on comprehension of print versus online information. *Technical Communication*, *53*, 416-426.

https://www.hcde.washington.edu/files/people/docs/HeadingsPrintOnline.pdf

Ben-Yehudah, G., & Brann, A. (2019). Pay attention to digital text: The impact of the media on text comprehension and self-monitoring in higher-education students with ADHD. *Research in Developmental Disabilities, 89*, 120-129. https://doi.org/10.1016/j.ridd.2019.04.001

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014, February). The influence of text annotation tools on print and digital reading comprehension. In Y. Eshet, A. Caspi, N. Geri, Y. Kalman, V. Silber-Varod, & Y. Yair (Eds.), *Proceedings of the 9th Chais Conference for Innovation in Learning Technologies* (pp. 28–35). Raanana, Israel: Open University Press.

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*, 206-219. https://doi.org/10.1016/j.cedpsych.2014.05.003

Butterfuss, R., & Kendeou, P. (2018). The role of executive functions in reading comprehension. *Educational Psychology Review*, *30*, 801-826. https://doi.org/10.1007/s10648-017-9422-6

Carr, N. (2010). *The shallows: How the internet is changing the way we think, read and remember*. London: Atlantic Books Ltd.

Carrier, L. M., Rosen, L. D., Cheever, N. A., & Lim, A. F. (2015). Causes, effects, and practicalities of everyday multitasking. *Developmental Review, 35*, 64-78. https://doi.org/10.1016/j.dr.2014.12.005

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 161-238). Mahwah, NJ: Erlbaum.

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, *42*, 288-325. https://doi.org/10.1111/1467-9817.12269

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coscarelli, C. V., & Coiro, J. (2015). Reading multiple sources online. *Revista Linguagem & Ensino*, *17*, 751-776. https://doi.org/10.15210/RLE.V17I3.15313

Daniel, D. B., & Woody, W. D. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education*, *62*, 18-23. https://doi.org/10.1016/j.compedu.2012.10.016

DeStefano, D., & LeFevre, J. A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, *23*, 1616-1641. https://doi.org/10.1016/j.chb.2005.08.012

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, *35*, 1297-1326. https://doi.org/10.1080/00140139208967394

Dobler, E. (2015). e-Textbooks: A Personalized Learning Experience or a Digital Distraction?. *Journal of Adolescent & Adult Literacy*, *58*, 482-491. https://doi.org/10.1002/jaal.391

Enea, S., Maresh, M., Miller, C., Pritzl, T., & Trieglaff, K. (2019). Impact of Online Reading on Mean Arterial Blood Pressure and Heart Rate. *Journal of Advanced Student Science, Spring 2019.* http://jass.neuro.wisc.edu/2019/01/301_3.pdf

E-READ, 2019. Stavanger Declaration. http://ereadcost.eu/wp-content/uploads/2019/01/StavangerDeclaration.pdf

Fishbein, B., Martin, M. O., Mullis, I. V., & Foy, P. (2018). The TIMSS 2019 Item Equivalence Study: examining mode effects for computer-based assessment and implications for measuring trends. *Large-scale Assessments in Education, 6*: 11. https://doi.org/10.1186/s40536-018-0064-z

Greenfield, P. M. (2009). Technology and informal education: What is taught, what is learned. *Science*, *323*, 69-71. https://doi.org/10.1126/science.1167190

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*, 93-103. https://doi.org/10.3758/MC.36.1.93

Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse Processes*, *37*, 187-204. https://doi.org/10.1207/s15326950dp3703_1

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Oxon: Routledge.

Hervet, G., Guérard, K., Tremblay, S., & Chtourou, M. S. (2011). Is banner blindness genuine? Eye tracking internet text advertising. *Applied Cognitive Psychology, 25*, 708-716. https://doi.org/10.1002/acp.1742

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *Journal of Technology, Learning and Assessment, 3*(4) https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1657

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of*

Educational Psychology, 94, 44-55. https://doi.org/10.1037/0022-0663.94.1.44

Jerrim, J., Micklewright, J., Heine, J. H., Salzer, C., & McKeown, C. (2018). PISA 2015: how big is the 'mode effect'and what has been done about it? *Oxford Review of Education, 44*, 476-493. https://doi.org/10.1080/03054985.2018.1430025

Ji, S. W., Michaels, S., & Waterman, D. (2014). Print vs. electronic readings in college courses: Cost-efficiency and perceived learning. *The Internet and Higher Education*, *21*, 17-24. https://doi.org/10.1016/j.iheduc.2013.10.004

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*, 524-532. https://doi.org/10.1177/0956797611430953

Kaakinen, J. K., Ballenghein, U., Tissier, G., & Baccino, T. (2018). Fluctuation in cognitive engagement during reading: Evidence from concurrent recordings of postural and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1671-1677. https://doi.org/10.1037/xlm0000539

Kaufman, G., & Flanagan, M. (2016). High-low split: Divergent cognitive construal levels triggered by digital and non-digital platforms. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2773-2777. https://doi.org/10.1007/s11145-003-8128-y

Kingston, N.M. (2008). Comparability of computer- and paper- administered multiple-choice tests for K–12 populations: A synthesis. *Applied Measurement in Education*, *22*, 22-37. https://doi.org/10.1080/08957340802558326

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, *95*, 163-182.

https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W (1998). *Comprehension. A paradigm for cognition*. Cambridge: Cambridge University Press.

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, *67*, 135-142. https://doi.org/10.1016/j.tate.2017.06.001

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education, 123*, 138-149. https://doi.org/10.1016/j.compedu.2018.05.005

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PLOS ONE, 8*, e56178. https://doi.org/10.1371/journal.pone.0056178

Lansink, J. M., & Richards, J. E. (1997). Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. *Child Development*, *68*, 610-620. https://doi.org/10.1111/j.1467-8624.1997.tb04224.x

Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology*, *59*, 101797. https://doi.org/10.1016/j.cedpsych.2019.101797

Latini, N., Bråten, I., & Salmerón, L. (2019). It's All about Sex: Reading Medium Affects Processing and Integration of Textual and Pictorial Information about

Human Sexuality. Manuscript submitted for publication.

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior*, *35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Lee, H.Y., & List, A. (2019). Processing of texts and videos: A strategy-focused analysis. *Journal of Computer Assisted Learning, 35*, 268-282. https://doi.org/10.1111/jcal.12328

Levine, L. E., Waite, B. M., & Bowman, L. L. (2007). Electronic media use, reading, and academic distractibility in college youth. *CyberPsychology & Behavior*, *10*, 560-566. https://doi.org/10.1089/cpb.2007.9990

Liew, C. L., Foo, S., & Chennupati, K. R. (2000). A study of graduate student end-users' use and perception of electronic journals. *Online Information Review*, *24*, 302-315. https://doi.org/10.1108/14684520010350650

List, A. (2018). Strategies for comprehending and integrating texts and videos. *Learning and Instruction, 57*, 34-46. https://doi.org/10.1016/j.learninstruc.2018.01.008

List, A., & Ballenger, E.E. (2019). Comprehension across mediums: the case of text and video. *Journal of Computing in Higher Education, 31*, 514-535. https://doi.org/10.1007/s12528-018-09204-9

Luyten, H., Merrell, C., & Tymms, P. (2017). The contribution of schooling to learning gains of pupils in Years 1 to 6. *School Effectiveness and School Improvement, Vol. 28*, 374–405. https://doi.org/10.1080/09243453.2017.1297312

Macedo-Rouet, M., Salmerón, L., Ros, C., Pérez, A., Stadtler, M., & Rouet, J. F. (2019). Are frequent users of social network sites good information evaluators?

An investigation of adolescents' sourcing abilities. *Journal for the Study of Education and Development, 43,* 1-17. https://doi.org/10.1080/02103702.2019.1690849

Mangen, A., & van der Weel, A. (2016). The evolution of reading in the age of digitisation: an integrative framework for reading research. *Literacy, 50*, 116-124. https://doi.org/10.1111/lit.12086

Margaryan, A., Littlejohn, A., & Vojt, G. (2011). Are digital natives a myth or reality? University students' use of digital technologies. *Computers & Education*, *56*, 429-440. https://doi.org/10.1016/j.compedu.2010.09.004

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*, 125-139. https://doi.org/10.1016/S0959-4752(02)00016-6

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43–72). New York, NY: Cambridge University Press.

Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., Glennerster, R., Green, P., Humphreys, M., Imbens, G., Laitin, D., Madon, T., Nelson, L., Nosek, B. A., Petersen, M., Sedlmayr R., Simmons, J. P., Simonsohn, U., van der Lahn, M. (2014). Promoting transparency in social science research. *Science*, *343*, 30-31. https://doi.org/10.1126/science.1245317

Mizrachi, D., Salaz, A. M., Kurbanoglu, S., & Boustany, J., on behalf of the ARFIS Research Group. (2018). Academic reading format preferences and behaviors among university students worldwide: A comparative survey analysis. *PloS ONE*, *13*, e0197444. https://doi.org/10.1371/journal.pone.0197444

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*, 229-241. https://doi.org/10.1016/j.cedpsych.2009.12.001

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., Simonsohn, U., Wagenmakers, E., Ware J. J., & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*: 0021. https://doi.org/10.1038/s41562-016-0021

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media?. *Computers in Human Behavior*, *54*, 301-309. https://doi.org/10.1016/j.chb.2015.07.043

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*, 1422-1425. https://doi.org/10.1126/science.aab2374

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*, 15583-15587. https://doi.org/10.1073/pnas.0903620106

Palfrey, J. G., & Gasser, U. (2011). *Born digital: Understanding the first generation of digital natives*. New York, NY: Basic Books.

Pashler, H., & Wagenmakers, E. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on*

Psychological Science, 7, 528-530. https://doi.org/10.1177/1745691612465253

Petko, D., Cantieni, A., & Prasse, D. (2017). Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *Journal of Educational Computing Research*, *54*, 1070-1091. https://doi.org/10.1177/0735633116649373

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *The Journal of Technology, Learning, and Assessment, 2*. n6 https://eric.ed.gov/?id=EJ905028.

Porges, S. W., & Raskin, D. C. (1969). Respiratory and heart rate components of attention. *Journal of Experimental Psychology, 81*, 497-503. https://doi.org/10.1037/h0027921

Prensky, M. (2001). Digital natives, digital immigrants, part 1. *On the Horizon*, 9(5), 1-6. https://doi.org/10.1108/10748120110424816

Prensky, M. (2009). H. sapiens digital: From digital immigrants and digital natives to digital wisdom. *Innovate: Journal of Online Education*, *5*(3): 1. https://www.learntechlib.org/p/104264/.

Randall, J. G., Beier, M. E., & Villado, A. J. (2019). Multiple routes to mind wandering: Predicting mind wandering with resource theories. *Consciousness and Cognition*, *67*, 26-43. https://doi.org/10.1016/j.concog.2018.11.006

Rideout, V. J. (2015). *The common sense census: Media use by tweens and teens*. Common Sense Media Incorporated. https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2015

Rose, E. (2011). The phenomenology of on-screen reading: University students' lived experience of digitised text. *British Journal of Educational Technology*, *42*, 515-526. https://doi.org/10.1111/j.1467-8535.2009.01043.x

Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, *52*, 200-215. https://doi.org/10.1080/00461520.2017.1329015

Rowsell, J. (2014). Toward a phenomenology of contemporary reading. *Australian Journal of Language and Literacy, 37*, 117-127. https://search.informit.com.au/documentSummary;dn=317701976558175;res=IELHSS

Salmerón, L., Sampietro, A., & Delgado, P. (2020). Using Internet videos to learn about controversies: Evaluation and integration of multiple and multimodal documents by primary school students. *Computers & Education, 148*, 103796. https://doi.org/10.1016/j.compedu.2019.103796

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS ONE*, *8*, e54402. https://doi.org/10.1371/journal.pone.0054402

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, *51*, 61-73. https://doi.org/10.1016/j.learninstruc.2017.01.002

Singer, L. M., & Alexander, P. A. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research*, *87*, 1007-1041. https://doi.org/10.3102/0034654317722961

Snow, C. E. (2002). Reading for understanding: Toward a research and development program in reading comprehension. Santa Monica, CA: RAND

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS ONE*, *8*, e54402. https://doi.org/10.1371/journal.pone.0054402

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*, 437-475. https://doi.org/10.1023/A:1009084430926

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, *85*, 698-739. https://doi.org/10.3102/0034654314566989

Thiede, K. W., Dunlosky, J., Griffin, T. D., & Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1267-1280. https://doi.org/10.1037/0278-7393.31.6.1267

van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. Magliano, & G. Schraw (Eds.), *Relevance instructions and goal- focusing in text learning* (pp. 123-140). Greenwich, CT: Information Age Publishing.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*, 360-372. https://doi.org/10.1080/0163853X.2017.1306677

van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior, 53*, 204-215. https://doi.org/10.1016/j.chb.2015.06.035

Wang, S., Jiao, H., Young, M.J., Brooks, T., & Olson, O. (2007). Comparability of computer-based and paper-and-pencil testing in K–12 reading assessments. A meta-analysis of testing mode effects. *Educational and Psychological Measurement, 68*, 5-24. https://doi.org/10.1177/0013164407305592

Wolf, M. (2018). *Reader, come home: The reading brain in a digital world*. New York, NY: HarperCollins

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*, 681-695. https://doi.org/10.3758/s13421-016-0589-8

[This page intentionally left blank]

Appendix

Appendix A

Additional content from Study 1

A1. References included in the meta-analysis

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning. *Journal of Experimental Psychology: Applied, 17*, 18-32. https://doi.org/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*, 1816-1828. https://doi.org/10.1016/j.chb.2012.04.023

Aydemir, Z., Öztürk, E., & Horzum, M. B. (2013). The effect of reading from screen on the 5th grade elementary students' level of reading comprehension on informative and narrative type of texts. *Kuram Ve Uygulamada Egitim Bilimleri, 13*, 2272-2276. https://doi.org/10.12738/estp.2013.4.1294

Baker, R. D. (2010). *Comparing the readability of text displays on paper, e-book readers, and small screen devices* (Doctoral dissertation). https://search.proquest.com/openview/1bfe9d4c681f59fddc1860217b565037/1?pq-origsite=gscholar&cbl=18750&diss=y

Bansi, S., Oudega, M., Koornneef, A., & van den Broek, P. (2016). *The influence of presentation medium and induced beliefs on reading comprehension: An eye-tracking study*. Poster session presented at the Scandinavian Workshop on Applied Eye Tracking 2016, Turku, Finland.

Bartell, A. L., Schultz, L. D., & Spyridakis, J. H. (2006). The effect of heading frequency on comprehension of print versus online information. *Technical Communication, 53*, 416-426. http://faculty.washington.edu/jansp/Publications/HeadingsPrintOnline.pdf

Beach, K. L. (2008). *The effect of media, text length, and reading rates on college student reading comprehension levels* (Doctoral dissertation). https://search.proquest.com/docview/89266558

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014). The influence of text annotation tools on print and digital reading comprehension. *Proceedings of the 9th Chais Conference for Innovation in Learning Technologies,* 28-35. http://innovation.openu.ac.il/chais2014/download/B2-2.pdf

Burkley, A. S. (2013). *The effect of liquid crystal displays on reading comprehension* (Doctoral dissertation). *https://search.proquest.com/openview/ac8aba8b5933dc6fb7a4f1e1384a9264/1?pq-origsite=gscholar&cbl=18750&diss=y*

Chen, D. (2015). *Metacognitive prompts and the paper vs. screen debate: How both factors influence reading behavior* (Master thesis). https://smartech.gatech.edu/handle/1853/53840

Chen, G., Cheng, W., Chang, T., Zheng, X., & Huang, R. (2014). A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education, 1*, 213-225. https://doi.org/10.1007/s40692-014-0012-z

Connell, C., Bayliss, L., & Farmer, W. (2012). Effects of eBook readers and tablet computers on reading comprehension. *International Journal of Instructional*

Media, 39, 131-141. https://www.questia.com/library/journal/1G1-289721760/effects-of-ebook-readers-and-tablet-computers-on-reading

Daniel, D. B., & Woody, W. D. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education, 62*, 18-23. https://doi.org/10.1016/j.compedu.2012.10.016

Delgado, P., & Salmerón, L. (2017). *Reading on print or on tablet: An eye-tracking study*. Poster session presented at the EARLI 2017 Conference*,* Tampere, Finland.

Duran, E. (2013). Efficiency in reading comprehension: A comparison of students' competency in reading printed and digital texts. *Educational Research and Reviews*, *8*, 258-269. https://doi.org/10.5897/ERR12.180

Eyre, J., Berg, M., Mazengarb, J., & Lawes, E. (2017). *Mode equivalency in PAT: Reading comprehension.* Wellington: NZCER. http://www.nzcer.org.nz/system/files/PAT%20Modes_report.pdf

Green, T. D., Perera, R. A., Dance, L. A., & Myers, E. A. (2010). Impact of presentation mode on recall of written text and numerical information: Hard copy versus electronic. *North American Journal of Psychology, 12*, 233-242. https://psycnet.apa.org/record/2011-04670-004

Grimshaw, S., Dungworth, N., McKnight, C., & Morris, A. (2007). Electronic books: Children's reading and comprehension. *British Journal of Educational Technology*, *38*, 583-599. https://doi.org/0.1111/j.1467-8535.2006.00640.x

Heij, M., & van der Meij, H. (2014). *The (in)effectiveness of PDF reading* (Bachelor thesis). http://essay.utwente.nl/65734/1/Heij%2C%20M.L.%20-%20s1189026%20%28verslag%29.pdf

Hermena, E. W., Sheen, M., AlJassmi, M., AlFalasi, K., AlMatroushi, M., & Jordan, T. R. (2017). Reading rate and comprehension for text presented on tablet and paper: Evidence from Arabic. *Frontiers in Psychology, 8*, 257. https://doi.org/10.3389/fpsyg.2017.00257

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *The Journal of Technology, Learning, and Assessment, 3*. https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1657

Hongler, K. I. (2015). *Superiority of paper as text presentation medium for effective and efficient learning – is it just an illusion?* (Master thesis). http://www.mmi-basel.ch/wp-content/uploads/2016/04/hongler_masterarbeit.pdf

Jeong, H. (2012). A comparison of the influence of electronic books and paper books on reading comprehension, eye fatigue, and perception. *The Electronic Library, 30*, 390-408. https://doi.org/10.1108/02640471211241663

Johnson, J. W. (2013). *A comparison study of the use of paper versus digital textbooks by undergraduate students* (Doctoral dissertation). http://scholars.indstate.edu/xmlui/bitstream/handle/10484/5376/Johnson,%20James%20Dissertation%20Final.pdf?sequence=2

Jones, M. Y., Pentecost, R., & Requena, G. (2005). Memory for advertising and information content: Comparing the printed page to the computer screen. *Psychology and Marketing, 22*, 623-648. https://doi.org/10.1002/mar.20077

Kaufman, G., & Flanagan, M. (2016). High-low split: Divergent cognitive construal levels triggered by digital and non-digital platforms. *Proceedings of the 2016 CHI*

Conference on Human Factors in Computing Systems, 2773-2777. https://doi.org/10.1145/2858036.2858550

Kerr, M., & Symons, S. (2006). Computerized presentation of text: Effects on children's reading of informational material. *Reading and Writing, 19*, 1-19. https://doi.org/10.1007/s11145-003-8128-y

Kim, D., & Huynh, H. (2008). Computer-based and paper-and-pencil administration mode effects on a statewide end-of-course english test. *Educational and Psychological Measurement, 68*, 554-570. https://doi.org/10.1177/0013164407310132

Kim, H. J., & Kim, J. (2013). Reading from an LCD monitor versus paper: Teenagers' reading performance. *International Journal of Research Studies in Educational Technology, 2*, 15-24. https://doi.org/10.5861/ijrset.2012.170

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *Plos One, 8*, e56178. https://doi.org/10.1371/journal.pone.0056178

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455-463. https://doi.org/10.1016/j.chb.2014.02.046

Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse Processes, 54*, 427-445. https://doi.org/10.1080/0163853X.2017.1319653

Liang, T., & Huang, Y. (2014). An investigation of reading rate patterns and retrieval outcomes of elementary school students with E-books. *Journal of Educational Technology & Society, 17*, 218. https://search.proquest.com/docview/1502989176

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Margolin, S. J., Driscoll, C., Toland, M. J., & Kegler, J. L. (2013). E-readers, computer screens, or paper: Does reading comprehension change across media platforms? *Applied Cognitive Psychology, 27*, 512-519. https://doi.org/10.1002/acp.2930

Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*, 367-378. https://doi.org/10.1016/S0169-8141(01)00043-9

McCrea-Andrews, H. J. (2014). *A comparison of adolescents' digital and print reading experiences: Does mode matter?* (Doctoral dissertation). http://tigerprints.clemson.edu/cgi/viewcontent.cgi?article=2337&context=all_dissertations

Morineau, T., Blanche, C., Tobin, L., & Guéguen, N. (2005). The emergence of the contextual role of the e-book in cognitive processes through an ecological and functional analysis. *International Journal of Human - Computer Studies, 62*, 329-348. https://doi.org/10.1016/j.ijhcs.2004.10.002

Niccoli, A. (2015). Paper or tablet reading recall and comprehension. *EDUCAUSE Review Online, 50.* https://er.educause.edu/articles/2015/9/paper-or-tablet-reading-recall-and-comprehension

Nishizaki, D. M. (2015). *The effects of tablets on learning: Does studying from a tablet computer affect student learning differently across educational levels* (Senior thesis). http://scholarship.claremont.edu/cmc_theses/1011/

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media? *Computers in Human Behavior, 54*, 301-309. https://doi.org/10.1016/j.chb.2015.07.043

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *Journal of Technology, Learning, and Assessment, 2*(6). https://eric.ed.gov/?id=EJ905028

Pomplun, M., Frey, S., & Becker, D. F. (2002). The score equivalence of paper-and-pencil and computerized versions of a speeded test of reading comprehension. *Educational and Psychological Measurement, 62*, 337-354. https://doi.org/10.1177/0013164402062002009

Porion, A., Aparicio, X., Megalakaki, O., Robert, A., & Baccino, T. (2016). The impact of paper-based versus computerized presentation on text comprehension and memorization. *Computers in Human Behavior, 54*, 569-576. https://doi.org/10.1016/j.chb.2015.08.002

Puhan, G., Boughton, K. A., & Kim, S. (2005). Evaluating the comparability of Paper-and-Pencil and computerized versions of a Large-Scale certification test. *ETS Research Report Series, 2005*(2), 1-15. https://doi.org/10.1002/j.2333-8504.2005.tb01998.x

Rasmusson, M. (2015). Reading paper – reading screen. *Nordic Studies in Education, 35*, 3-19. https://www.idunn.no/np/2015/01/reading_paper__reading_screen_-_a_comparison_of_reading_l

Sackstein, S., Spark, L., & Jenkins, A. (2015). Are e-books effective tools for learning? reading speed and comprehension: iPad vs. paper. *South African Journal of Education, 35*(4), 1-14. https://doi.org/10.15700/saje.v35n4a1202

Seehafer, H. (2014). Effects of learning style on paper versus computer based reading comprehension. *The Red River Psychology Journal, 2014*(1). https://www.mnstate.edu/RRpsychjournal/

Simian, M., Malbrán, C., Fonseca, L., Grinberg, S., Gattas, M. J., Tascón, R., & Cirigliano, S. (2016). *Reading narrative, expository and discontinuous texts on paper versus screen: Impact on reading comprehension in Argentine high school kids*. Poster session presented at 33rd annual meeting of the Society for the Scientific Studies of Reading (SSSR), Porto, Portugal.

Singer, L. M., & Alexander, P. A. (2017). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education, 85*, 155-172. https://doi.org/10.1080/00220973.2016.1143794

Taylor, A. K. (2011). Students learn equally well from digital as from paperbound texts. *Teaching of Psychology, 38*, 278-281. https://doi.org/10.1177/0098628311421330

Thompkins, P., Baker, L., & DeWeyngaert, L. (2013). *Paper or pixel? The influence of text format and metacognition on student reading comprehension*. Communication presented at the AERA 2016 Conference, Washington, DC, USA.

Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior, 21*, 377-394. https://doi.org/10.1016/j.chb.2004.02.007

Wells, C. L. (2012). *Do students using electronic books display different reading comprehension and motivation levels than students using traditional print books?* (Doctoral dissertation). https://search.proquest.com/docview/1266822858

A2. Coding Manual

General guidelines:

The purpose of this guide is to describe the items in the spreadsheet in which studies are codified and to guide the coding process. If you have a doubt when codifying some information, please do not overlook it and annotate it by adding a comment to the corresponding cell. Indicate the page of the paper where the information related to your doubt can be find.

As you can see, sections are separated by colours. If any paper has more than one experiments or independent variables (besides the reading medium variable), then they have to be coded in a different line (as it is explained later). In such case, please add new lines by duplicating the first one.

Please note that many cells have a drop-down list to fill it out.

If the information needed to fill any cell are not reported by any paper, then annotate "N/r" (or select it from the drop-down list).

If the information needed to fill any cell is not applicable for any paper because of its characteristics, then annotate "N/a" (or select it from the drop-down list). In this case, please insert a comment in the cell to indicate the reason. For example, if any study has a within-subject design, the information related to the group allocation (see section "D-Experimental Design" below) would be "N/a".

If you consider any information or clarification that is important to a particular data in a particular cell, you can also insert a comment.

Section-by-section guidelines:

A – ID

The purpose of this section is to assign an ID number to each paper.

Some papers report more than one experiment or conditions. In that case, as you already read, you should create (by duplicating) a line for each experiment or each condition (i.e., additional independent variables besides the medium condition), respectively. The ID will be encoded by adding a second number to the paper ID, for instance: 1-1, 1-2, 1-3, 1-4, …

The additional independent variables (remember: one per line) will be coded in section "H-Additional Independent Variable" (see below)

B – Citation Parameters

In this section, some general factors related to publishing features are codified: Author/s, Year of publication, Source name, Title, Country of study, Publication language, Publication type and

Keywords. You should also add a link to the paper in the last subsection (papers will be uploaded before to Google Drive)

If any paper has more than one experiment or conditions within the same experiment, then this information will be repeated in each line.

C – Participants

In this section, sample features are codified: Sample Size, Age, Number of males and females, Educational level, Presence of participants with learning disabilities and Presence of participants with atypical development.

You will find a subsection for the whole sample and for each group, corresponding to cells for size, age and gender proportion.

*Sample size: annotate the initial sample size (observed N will be annotated in section "K-Results")

*Sampling Method:

- **Probability**: A probability sampling method utilizes some form of random selection, by setting up some process or procedure that assures that the different units in population have equal probabilities of being chosen.

- **Nonprobability**: sampling does not involve random selection of the population sample.

- N/r

*Same sample?: Fill in this cell only in case that the same paper has more than one line in the spreadsheet (i.e., more than one experiment or condition), to indicate if the sample is or not the same.

*Educational levels: You will find several possibilities in the drop-down list:

- Pre-school,

- Different school grades (1 to 12)

- Undergraduates

- Graduates or professionals.

In case they specify different grades or levels, do select "More than one grade or level" and specify them by adding a comment.

*Special population: Select YES the whole sample, YES a subgroup or NO. If YES, specify why they are special by adding a comment (e.g., learning disabilities, L2 readers…)

*n: annotate de initial subsamples size (observed n will be annotated in section "K-Results")

*Number of males or females: You will find four columns (number and percentage of males and the same for females). You do not have to fill in all but just the ones that are reported by the paper (Sometimes the exact number of males or females is reported, while other times they report the percentage of one of the genders)

* Age: You will find different columns that you can use depending on what information is reported by the paper:

- Mean and standard deviation.
- Median and/or the range.
- Use "/" to separate different data from several age subgroups within the on-paper or on-screen groups.

*Educational level: You will find a drop-down list with standard educational levels for pre-, primary and secondary education. In case that any paper reports educational levels from a different system please select the one for the corresponding age, as follows:

- Pre-school: 3-6 years-old
- Grade 1: 6-7
- Grade 2: 7-8
- Grade 3: 8-9
- Grade 4: 9-10
- Grade 5: 10-11
- Grade 6: 11-12
- Grade 7: 12-13
- Grade 8: 13-14
- Grade 9: 14-15
- Grade 10: 15-16
- Grade 11: 16-17
- Grade 12: 17-18

D – Experimental Design

In this section, the following features are codified:

- Type of design: Select from the drop-down list if they used a within- or a between-subjects design.

- Group allocation: How participants were assigned to groups. Select from the list, following these clarifications:

- **Random**:

A. randomly after matching, yoking, stratification, blocking, etc. The entire sample is matched or blocked first, then assigned to treatment and comparison groups within pairs or blocks. This does not refer to blocking after treatment for the data analysis.
B. randomly without matching, etc. This also includes cases when every other person goes to the control group.

- **Quasi-random**: procedure presumed to produce comparable groups (no obvious differences). This applies to groups which have individuals apparently randomly assigned by some naturally occurring process, e.g. next person to walk in the door. The key here is that the procedure used to select groups doesn't involve individual characteristics of persons so that the groups generated should be essentially equivalent.

- **Non-random but matched or controlled** (in this case, please annotate which variables are matched or controlled by adding a comment): Matching refers to the process by which comparison groups are generated by identifying individuals or groups that are comparable to the treatment group using various characteristics of the treatment group. Statistical control refers to inclusion of the matching variable as a covariate in an ANCOVA or multiple regression analysis. Matching can be done individually, e.g., by selecting a control subject for each intervention subject who is the same age, gender, and so forth, or on a group basis, e.g., by selecting comparison schools that have the same demographic makeup and academic profile of treatment schools. Similarly, statistical control variables can be used at either the individual or school level.

- **Non-random not controlled**.

- N/r (because within-subject design)

- N/r

- Setting: Select the type of room where the experiment was applied (i.e., room, computer lab, lab, home, classroom, etc.)

E – Digital Device

Select the type of digital device that was used in the experiment:

- Desktop
- Laptop
- Computer (in case that kind of computer is not specified)
- Tablet
- E-reader
- Smartphone

Additionally:

- Annotate what kind of display was used (e.g. LCD, TFT, …) in case they reported it.

- Annotate the screen size.

*As it is indicated in the spreadsheet, you can fill out additional columns in case that any paper used different types of devices. But please note that if different devices are used as different experimental conditions, they must be codified in different lines (not in additional columns).

F – Experimental Material

In this section, the following features are codified:

- Text Format: Select whether they use text, text with images, with hypertext or with both. Here you can add more "Text Format" subsections in case there are 2 or more conditions for this.

- Type of text: Select the type of text/s:

- Narrative

- Descriptive
- Expository
- Argumentative
- Directive.
- Poetry.

(In the same way, you can also add more subsections for this).

- Number of texts: annotate how many texts they used.
- Word count: Use “XXX-YYY” when they report a range of number of words. Use “/” to indicate different numbers of words for different texts.
- Typography: Annotate the typography features reported, e.g., font, size, colour, spacing, etc.
- Text/s Language.
- Scroll?: Please indicate (by drop-down list) if top-down scrolling is needed to read digital texts.

*As it is indicated in the spreadsheet, you can fill out additional columns in case that any paper used different text formats or types of texts. But please note that if they are used as different experimental conditions, they must be codified in different lines (not in additional columns).

G – Comprehension Task

Here, the following features are codified:

- Total length: record total length (time) in case they report the length of the whole experiment.
- Practice with digital tools?: Do indicate (by drop-down list: YES or NO) if participants practiced with digital tools before using them in task/s.
- Official assessment of academic achievement?: Do indicate (by drop-down list: YES or NO) if data were collected from an official assessment of academic achievement (e.g., PISA)
- Comprehension type: Select form the drop-down list if they asses:
  - Textual: the task asks for textual content, details or shallow level comprehension.
  - Inferential: the task asks for high level comprehension. It demands inferences between different parts of the texts or between text and previous knowledge.
  - Textual + Inferential: same task use both types of questions.
  - Other: please do specify which type of comprehension is demanded by adding a comment to the cell
  - N/r
- Task type: Select the type of task they used to asses de reading comprehension:
  - Standardized test
  - Multiple choice

- Short answers
- Cloze task
- Summarization
- Error detections
- Other
- N/r

*In case you select “Standardized test”, please, do specify, by adding a comment to the cell, what is the tool they used. In case you select “Other”, do specify what task by adding a comment.

- Number of items: Record the number of items that constitute the task.

- Allotted reading time: Select “free” or “limited” from the drop-down list.

- Length of each particular task: Annotate it in minutes.

- Back to texts?: Do indicate (by drop-down list: YES or NO) if participants are allowed to back to texts once they have read the questions or performing the required task.

- Reliability: Annotate the task reliability.

- Answering medium: Select if participants responded:

  - by writing it down

  - by using digital devices

  - by using the same medium as they read in each condition

  - orally.

- Instructions: Please copy/paste literal instructions that was given to participants.

- Depth of processing?: Do indicate (by drop-down list: YES or NO) if task instructions promote depth of processing. Do specify how they did it by adding a comment (e.g., note taking, highlighting, selecting key words…)

* You will find three subsections to fill out just in case they report two different tasks. If there are more than two tasks, please add additional subsections by duplicating it.

<u>H – Dependent Variables</u>

Mark with an “x” to indicate what dependent variable/s are included in the experimental design. You will find different columns for several possible independent variables, respectively. If any paper reports a dependent variable that is not included, annotate it (specify the variable) in the column “Other”.

* Clarifications:

- “Learning”: participants are required to study the materials (not only to read and comprehend), and their learning is evaluated.

- "PoP" (Prediction of Performance): participants make an estimation of their success in the later responses before reading the questions.

- "Confidence judgement": participants make an estimation of their performance after they respond.

- "Calibration": they assess *calibration*, i.e., the absolute degree to which a person's perception of performance fits with his or her actual performance.

- "Resolution": they assess *resolution*, i.e., the relative degree to which a person's judgments can predict the likelihood of correct performance of one item relative to another (Nelson, 1996)

I – Controlled Variables

Mark each controlled variable by selecting "Design" or "Analysis" from the drop-down list to indicate, at the same time, whether they were controlled in the experimental design or in the statistical analysis. You will find a list of several possible confounders. If any paper reports a controlled variable that is not included in the spreadsheet, annotate it (specify the variable by adding a comment) in the column "Other".

At the end of this subsection, you will find several columns dedicated to possible confounders that should be controlled in case of a within-participants design. In this case, given that these variables must be controlled in the experimental design, just mark with an "X" to indicate which of them is controlled.

* Clarifications:

- "Texts" refers to they used the same texts in different medium conditions.

- "Text formatting" refers to they used the same text features related to typography, paragraph formatting, images etc.

J – Additional Independent Variables

If any paper asses 2 or more independent variables (by means of different experiments or conditions), in this section you can select the independent variable for each particular excel line for the same paper (Remember: you should add additional lines for each independent variable or each experiment). You will find a drop-down list of several possible independent variables. If any paper reports an independent variable that is not included, select other and specify the variable by adding a comment.

* Clarification:

- "Depth of Processing" refers to a guided-by-instructions reading (by asking participants for activities such summarize, draw the main ideas or keywords, note-taking, etc.). Please notice that, in this case, these activities wouldn't be the required answer but additional activities before participants answer the task.

In the second column, you must specify the value for the independent variable for each condition (remember one condition per line)

K - Results

* Please notice that, as we are using different lines in case of different conditions for additional factors, only simple effects should be codified. In case they report only the main effect, insert data and please do specify this circumstance in the column "Additional Info"

You will find 5 subsections that allowed you to codify:

- Three different reading comprehension measures ("Reading Comp. 1, 2 & 3)". You can use "Reading Comp. 2 (and 3)"in case they assessed reading comprehension by means of more than 1 task and they report the results separately. If there are more than three different scores for reading comprehension you can add more subsections by duplicating them.
- Reading time or reading speed.
- Reading efficiency, in case they report comparisons related to the ratio of invested time and comprehension scores.

Within each dependent variable subsection, you will find:

- At first, a column in which you must to specify the measure they used for each dependent variable.
- Observed N: Annotate the number of subjects who were actually measured.
- Descriptive statistics:
    - o Scores type: select from the drop-down list if they report total scores or percentage of correct answers.
    - o Mean, SD, Variance (σ^2) and/or SE: annotate it differentiating between paper and screen (as you will find in the spreadsheet)
- Effect size: Select from the list the statistic they used (in case you select other, please annotate the statistic by adding a comment) and annotate the value.
- Contrast statistic: Select from the list the statistic* they used (in case you select "other" in "Statistic" cell, please annotate the statistic they used by adding a comment); annotate the value and the *p*-value.

* F (1 d.f.) = analysis of variance of 2 groups (1 degree of freedom)

F (2+ d.f.) = analysis of variance of more than 2 groups (2 or more degrees of freedom)

t = independent t-test anaylisis

X^2 (1 d.f.) = Chi-square from 2x2 table (1 degree of freedom)

X^2 (2+ d.f.) = Chi-square from larger thah 2x2 table (2 or more degree of freedom)

- Significant differences but NO DATA: this is rare, but you could find some paper that reports significant differences without reporting any data. In this case marc with an "x" this cell.

- Favoured Group: Select which group was better. Remember that you cannot rely on simple numerical values to determine which group is favoured. For instance, if errors are measured, fewer errors are better than more, so in this case a lower number indicates a better performance.

Additional info: please record here some information that you consider relevant but is not requested.

Table A1

Descriptive characteristics of the variables coded for each reading media comparison, from the studies using a between-participants design.

Study/Comparison*	Publishing status	Sampling method	Group allocation	Sample size	Educational level	Text length	Testing medium	Digital device	Reading time frame	Text genre	Scroll	Type of test	Type of compre-hension	Open testing	Explicit strategic req.
Ackerman & Goldsmith, 2011 (Exp. 1)[2]	Yes	Non-probability	Random	70	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Mix	No	No
Ackerman & Goldsmith, 2011 (Exp. 2)[2]	Yes	Non-probability	Random	74	Undergraduates	Large	Same for reading	Computer	Free	Informational	Yes	R-C	Mix	No	No
Ackerman & Lauterman, 2012 (Exp. 1)a[2]	Yes	Non-probability	Random	41	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Mix	No	No
Ackerman & Lauterman, 2012 (Exp. 1)b[2]	Yes	Non-probability	Random	39	Undergraduates	Large	Same for reading	Computer	Free	Informational	Yes	R-C	Mix	No	No
Ackerman & Lauterman, 2012 (Exp. 2)[2]	Yes	Non-probability	Random	76	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Mix	No	No
Aydemir et al., 2013[2]	Yes	Non-probability	N/r	60	Grade 5	N/r	N/r	Computer	Free	Mix[8]	N/r	R-C	Mix	N/r	No
Bartell et al., 2006	Yes	Non-probability	Non-random	239	Undergraduates	Large	Same for reading	Computer	N/r	Informational	Yes	R-C	N/r	No	No
Beach, 2008a	No	Non-probability	Random	30	Undergraduates	Large	Same for reading	Computer	N/r	Informational	N/r	R-C	Mix	N/r	No
Beach, 2008b	No	Non-probability	Random	43	Undergraduates	Short	Same for reading	Computer	N/r	Informational	N/r	R-C	Mix	N/r	No
Ben-Yehudah & Eshet-Alkalai, 2014a	Yes	Non-probability	Random	46	Undergraduates	Short	Same for reading	Computer	Free	Informational	Yes	R-C	Mix	No	Yes
Ben-Yehudah & Eshet-Alkalai, 2014b	Yes	Non-probability	Random	47	Undergraduates	Short	Same for reading	Computer	Free	Informational	Yes	R-C	Mix	No	No
Burkley, 2013	No	Non-probability	Random	33	Undergraduates	N/r	Paper	Computer	N/r	N/r	Yes	Std.	Mix	Yes	No
Chen et al., 2014a	Yes	Non-probability	Random	45[5]	Undergraduates	Large	Paper	Computer	Limited	Informational	Yes	Std.	Mix	No	No
Chen et al., 2014b	Yes	Non-probability	Random	45[5]	Undergraduates	Large	Paper	Hand-held	Limited	Informational	No	Std.	Mix	No	No
Chen, 2015	No	Non-probability	N/r	92	Undergraduates	Large	Paper	Computer	Free	Informational	Yes	R-C	Textual	No	Yes
Connell et al., 2012a[2]	Yes	Non-probability	Random	104[5]	Undergraduates	Large[3]	Paper	Hand-held	Free	Informational	No[3]	Std.	Mix	No[3]	No
Connell et al., 2012b[2]	Yes	Non-probability	Random	98[5]	Undergraduates	Large[3]	Paper	Hand-held	Free	Informational	No[3]	Std.	Mix	No[3]	No
Daniel & Woody, 2013a	Yes	Non-probability	Random	59	Undergraduates	Large	N/r	Computer	Free	Informational	Yes	R-C	N/r	No	No
Duran, 2013[1]	Yes	Non-probability	Random	207	Undergraduates	N/r	N/r	Computer	N/r	Mix	N/r	R-C	N/r	N/r	No

Eyre et al., 2017a[2]	No	Non-probability[3]	Non-random	71818[3]	Grades 4 to 6	Short[3]	Same for reading	Computer	Limited[3]	Mix[3]	Yes[3]	Std.	Mix[3]	Yes[3]	No
Eyre et al., 2017b[2]	No	Non-probability[3]	Non-random	82759[3]	Grades 7 to 10	Short[3]	Same for reading	Computer	Limited[3]	Mix[3]	Yes[3]	Std.	Mix[3]	Yes[3]	No
Green et al., 2010	Yes	Non-probability	Random	54[6]	Undergraduates	N/R	Same for reading	Computer	Limited	Informational	Yes	R-C	Textual	No	No
Grimshaw et al., 2007a	Yes	Non-probability	Controlled	51	Elementary school	Large	N/r	Computer	N/r	Narrative	No	R-C	Mix	Yes	No
Grimshaw et al., 2007b	Yes	Non-probability	Controlled	55	Elementary school	Large	N/r	Computer	N/r	Narrative	No	R-C	Mix	Yes	No
Higgins et al., 2005a	Yes	Non-probability	Random	111[5]	Grade 4	Short[3]	Same for reading	Computer	Free	Mix[3]	No	Std.	Mix[3]	Yes	No
Higgins et al., 2005b	Yes	Non-probability	Random	108[5]	Grade 4	Short[3]	Same for reading	Computer	Free	Mix[3]	Yes	Std.	Mix[3]	Yes	No
Hongler, 2015a	No	Non-probability	Random	36[5]	Undergraduates	Large	Digital	Computer	Free	Informational	Yes	R-C	Mix	No	No
Hongler, 2015b	No	Non-probability	Random	36[5]	Undergraduates	Large	Digital	Hand-held	Free	Informational	Yes	R-C	Mix	No	No
Johnson, 2013	No	Non-probability	Random	233	Undergraduates	Large[3]	Paper	Hand-held	Free[3]	Informational	Yes[3]	R-C	Mix	Yes[3]	No
Jones et al., 2005	Yes	Non-probability	Random	48	Mix[4]	Short	Paper	Computer	N/r	Informational	No	R-C	Textual	No	No
Kaufman & Flanagan, 2016 (Study 2)	No	Non-probability	Random	81	Undergraduates	N/r	Paper	Computer	N/r	Narrative	N/r	R-C	Mix	N/r	No
Lauterman & Ackerman, 2014 (Exp. 1)[2]	Yes	Non-probability	Random	87	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Mix	No	No
Lauterman & Ackerman, 2014 (Exp. 2)[2]	Yes	Non-probability	Random	76	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Mix	No	Yes
Lenhard et al., 2017[2]	Yes	Probability	Random	2807	Grades 1 to 3	Short[3]	Same for reading	Computer	Limited	Mix	No	Std.	Mix	No[3]	No
Mangen et al., 2013[2]	Yes	Non-probability	Random[3]	72	Grade 10	Large	Same for reading	Computer	Limited	Mix	Yes	Std.	Mix	Yes	No
Margolin et al., 2013a	Yes	Non-probability	Random	45[5]	Undergraduates	Short	Paper	Computer	Free	Mix	Yes	R-C	Inferential	No	No
Margolin et al., 2013b	Yes	Non-probability	Random	45[5]	Undergraduates	Short	Paper	Hand-held	Inferential	No	No	R-C	Inferential	No	No
Mayes et al., 2001 (Exp. 1)	Yes	Non-probability	Random	40	Undergraduates	Large	Same for reading	Computer	Limited	Informational	Yes	R-C	Textual	N/r	No
McCrea-Andrews, 2014	No	Non-probability	Random	36	Grade 6	Large	N/r	Hand-held	Free	Narrative	Yes	Std	Mix	N/r	Yes
Morineau et al., 2005	Yes	Non-probability	Random	40	Graduates or prof.	N/r	Paper	Hand-held	Free[3]	Narrative	Yes[3]	R-C	Mix[3]	No	No
Niccoli, 2015	Yes	Non-probability	Random	231	Graduates or prof.	Short	Paper[3]	Hand-held	Free[3]	Informational	Yes[3]	R-C	Mix[3]	No[3]	No
Nishizaki, 2015 (Exp. 1)a[1]	No	Non-probability	Random	40	Grade 4	Short	Paper	Hand-held	Limited	Narrative	N/r	Std.	Mix	No	No

Nishizaki, 2015 (Exp. 1)b	No	Non-probability	Random	40	Undergraduates	Short	Paper	Hand-held	Limited	Informational	N/r	Std.	Mix	No	No
Nishizaki, 2015 (Exp. 2)	No	Non-probability	Random	80	Undergraduates	Short	Paper	Hand-held	Limited	Informational	N/r	Std.	Mix	No	No
Norman & Furnes, 2016 (Exp. 1)a	Yes	Non-probability	Random	37[5]	Undergraduates	Large	Paper	Computer	Limited	Informational	Yes	R-C	Textual	No	No
Norman & Furnes, 2016 (Exp. 1)b[7]	Yes	Non-probability	Random	63[5]	Undergraduates	Large	Paper	Hand-held	Limited	Informational	Yes	R-C	Textual	No	No
Norman & Furnes, 2016 (Exp. 2)	Yes	Non-probability	Random	50	Undergraduates	Large	Paper	Computer	Limited	Informational	Yes	R-C	Textual	No	No
Pommerich, 2004 (Exp. 1)	Yes	Non-probability	Random	1893	Grades 11 & 12	N/r	Same for reading	Computer	Limited	N/r	Yes	Std.	Mix	Yes	No
Pommerich, 2004 (Exp. 2)a	Yes	Non-probability	Random	2175	Grades 11 & 12	N/r	Same for reading	Computer	Limited	N/r	Yes	Std.	Mix	Yes	No
Pommerich, 2004 (Exp. 2)b	Yes	Non-probability	Random	2082	Grades 11 & 12	N/r	Same for reading	Computer	Limited	N/r	No	Std.	Mix	Yes	No
Porion et al., 2016	Yes	Non-probability	N/r	72	Grades 9 & 10	N/r	Paper	Computer	Limited	Informational	No	R-C	Mix	No	No
Puhan et al., 2005	Yes	Probability	Non-random	2224	Graduates or prof.	N/r	Same for reading	Computer	Limited	N/r	N/r	Std.	N/r	N/r	No
Seehafer, 2014	Yes	Non-probability	Random	67	Undergraduates	Short	Same for reading	Computer	N/r	Narrative	No	R-C	Inferential	N/r	No
Simian et al., 2016	No	Non-probability	N/r	87	Grade 8	Short	Paper[3]	Hand-held	Free[3]	Mix	Yes[3]	Std.	Mix	Yes	No
Taylor, 2011a[2]	Yes	Non-probability	Random	34	Undergraduates	Large	Paper	Computer[3]	Free[3]	Informational	Yes[3]	Std.	Textual[3]	No	No
Taylor, 2011b[2]	Yes	Non-probability	Random	35	Undergraduates	Large	Paper	Computer[3]	Free[3]	Informational	Yes[3]	Std.	Textual[3]	No	Yes
Wästlund et al., 2005	Yes	Non-probability	Controlled	76	Undergraduates	Large	Same for reading	Computer	Limited	Informational[3]	Yes[3]	Std.	Mix	Yes[3]	No
Wells, 2012	No	Non-probability	Random	152	Grades 6-12	Short	Same for reading	Hand-held	Limited	Mix	No	Std.	Mix	Yes	No

Note. *Letters after the publication year differentiate several comparisons from the same study. [1]Comparison excluded as it was identified as outlier. [2]The necessary statistical data was provided by authors following a personal request. [3]Information provided by authors following a personal request. [4]Sample composed by undergraduates and professionals with various educational levels. [5]Control group sample size was divided by two (see Method section). [6]Whole sample size was 82, but they were randomly assigned to three groups and only two groups participated in the reading media comparison (each group was considered as consisting of 27 participants). [7]Two comparisons with tablet and e-reader as digital device, respectively, were collapsed into this effect size. [8]Authors personally provided necessary statistical data only from narrative texts.

Table A2

Descriptive characteristics of the variables coded for each reading media comparison, from the studies that used a within-participants design.

Study/Comparison*	Publishing status	Sampling method	Sample size	Educational level	Text length	Testing medium	Digital device	Reading time frame	Text genre	Scroll	Type of test	Type of compre-hension	Open testing	Explicit strategic req.
Baker, 2010	No	Non-probabibily	100	Undergraduates	Short	Paper	Hand-held	Free	N/r[3]	No	Std.	Mix	N/r	No
Bansi et al., 2016[1]	No	Non-probabibily	29	Undergraduates	Short	N/r	Computer	Free	Informational	No	R-C	Mix	No	No
Delgado & Salmerón, 2017	No	Non-probabibily	69	Undergraduates	Short	Same for reading	Hand-held	Free	Informational	No	R-C	Mix	No	No
Heij & van der Meij, 2014	No	Non-probabibily	16	Undergraduates	Large	Paper	Computer	Free	Informational	Yes	R-C	Mix	No	No
Hermena et al., 2017	Yes	Non-probabibily	24	Undergraduates	Short	Orally	Hand-held	Free	Narrative	No	R-C	N/r	No	No
Jeong, 2012	Yes	Non-probability	56	Grade 6	Short	N/r	Computer	N/r	Narrative	Yes	R-C	Textual	N/r	No
Kerr & Symons, 2006	Yes	Non-probabibily	60	Grade 5	Short	N/r	Computer	Free	Informational	Yes	R-C	Mix	No	No
Kim & Huynh, 2008	Yes	Non-probabibily	439	Middle & High School	Short	Same for reading	Computer	Free	N/r	No	Std.	Inferential	Yes	No
Kim & Kim, 2013	Yes	Non-probabibily	108	Grade 11	Short[2]	Same for reading	Computer	Free[2]	Informational	Yes[2]	Std.	N/r	N/r	No
Kretzschmar et al., 2013a	Yes	Non-probabibily	35	Undergraduates	Short	Orally	Hand-held	Free	Mix	Yes	R-C	Textual	N/r	No
Kretzschmar et al., 2013b	Yes	Non-probabibily	21	Retired professionals	Short	Orally	Hand-held	Free	Mix	Yes	R-C	Textual	N/r	No
Liang & Huang, 2013	Yes	Non-probabibily	24	Grade 6	Large	Paper	Hand-held	Limited	Informational	Yes	R-C	Textual	No	No
Pomplun et al., 2002	Yes	Non-probabibily	215	Undergraduates	Short	Same for reading	Computer	Limited	Informational	Yes	Std.	Mix	No	No
Rasmusson, 2015	Yes	Non-probabibily	117	Grade 9	Short	Same for reading	Computer	Limited[2]	Mix	Mix[2]	Std.	Mix	Yes[2]	No
Sackstein et al., 2015a[4]	Yes	Non-probabibily	54	Grade 10	Short[2]	N/r	Hand-held	Free	Informational	No	Std.	Mix	No	No
Sackstein et al., 2015b	Yes	Non-probabibily	14	Undergraduates	Short[2]	N/r	Hand-held	Free	Informational	No	Std.	Mix	No	No
Singer & Alexander, 2017	Yes	Non-probabibily	90	Undergraduates	Short	Same for reading	Computer	Free[2]	Informational	No[2]	R-C	Mix	No	No
Thompkins et al., 2016	No	Non-probabibily	60	Undergraduates	Large	N/r	Computer	Limited	Informational	Yes[2]	R-C	Textual	No[2]	No

Note. *Letters in some references differentiate several comparisons from the same study. [1]The necessary statistical data was provided by authors following a personal request. [2]Information provided by authors following a personal request. [3]A selection of some texts from a standardized test was used, but texts genre is not specified. [4]Two different comparisons with Grade 10 students were collapsed into this effect size.

Appendix B

Materials used in Study 2

B1. The texts and the comprehension questions.

TEXT 1

Validación social

Según el fenómeno de la validación social, antes de decidir qué hacer en una situación observamos lo que otros individuos hacen o han hecho en la misma situación. Si son muchos los individuos que se han decidido por una determinada actuación, estamos más inclinados a adoptarla, porque percibimos esa idea más correcta, más válida.

El fenómeno de la validación social se ha estudiado mediante experimentos. Por ejemplo, en una invernal mañana de finales de los sesenta, un hombre se detuvo en una acera en medio del ajetreo neoyorquino y, durante 60 segundos, levantó su mirada hacia el cielo, sin mirar a nada en concreto. Ese hombre era un experimentador que trataba de estudiar el efecto ejercido por esa acción sobre los peatones.

Un 4 por ciento de los peatones se unió al experimentador que miraba al cielo, pero la mayoría simplemente se desvió o pasó de largo.

El experimento se repetiría luego con ligeras modificaciones, con las que un gran número de peatones fue inducido a detenerse y a mirar hacia arriba. En una primera modificación, al colocar cinco experimentadores mirando el cielo, no uno solo, se introdujo la influencia de la validación social en el experimento. El porcentaje de peatones que se agregó a los experimentadores fue del 18 por ciento, más del cuádruple del experimento inicial. Una segunda modificación incluyó un grupo de 15 experimentadores, que consiguió que el 40 por ciento de los peatones se uniera hasta casi detener el tráfico durante un minuto. Por tanto, los experimentos concluyeron que a mayores grupos iniciales de experimentadores, mayor número de peatones se veían afectados por la validación social.

Sin embargo, conclusiones menos obvias resultan cuando la validación social produce el efecto contrario del que se persigue. Un ejemplo lo constituye la comprensible, aunque mal orientada, tendencia de los educadores sanitarios a lanzar campañas informativas sobre un problema lamentando su excesiva frecuencia. Estas campañas recalcan que el uso de alcohol y drogas es elevado, que las tasas de suicidios entre adolescentes son alarmantes o que quienes contaminan degradan el medio. A pesar de que las campañas son tan ciertas como bienintencionadas, quienes las crean han olvidado algo básico sobre el proceso de validación social. Bajo la aseveración “Mira cuánta gente está haciendo algo tan

INDESEABLE" se esconde el poderoso y devastador mensaje de la validación social "Mira CUÁNTA GENTE ESTÁ haciendo algo tan indeseable". Como consecuencia, las investigaciones sobre el efecto negativo de la validación social en las campañas informativas demuestran que tienen un efecto bumerán, llegando incluso a estimular aquello que pretenden evitar.

QUESTIONS:

1. Siguiendo las conclusiones de los experimentos sobre el observador y los peatones, en qué campaña informativa es de esperar mayor influencia negativa de la validación social:

a) Campaña para prevenir el uso del móvil al volante.
b) Campaña informativa contra el uso de drogas duras.
c) Campaña para prevenir el abuso a menores.
d) Campaña informativa contra las centrales nucleares.

2. Qué eslogan te parece más efectivo para evitar el efecto negativo de la validación social en las campañas informativas para la prevención de drogas:

a) Los jóvenes sanos no consumen drogas.
b) La juventud y las drogas no hacen una buena combinación.
c) Que no te engañen, no consumas drogas.
d) Se diferente, diviértete sin drogas.

3. En qué se parecen el experimento del observador y las campañas informativas descritas en el texto:

a) En ambos casos aparece la validación social, pero no se parecen en las consecuencias de la misma.
b) En ambos casos se observa un fuerte efecto de la validación social empezando con poblaciones de referencia pequeñas.
c) En ambos casos los sujetos cambian su conducta tal y como se preveía por la validación social.
d) En ninguno de los casos los sujetos se influencian por la magnitud del grupo de referencia.

4. Según el fenómeno de la validación social, en una situación social:

a) Observamos lo que otros hacen o han hecho antes de decidir qué hacer.
b) Decidimos qué hacer de acuerdo a la norma establecida.
c) Decidimos qué hacer anticipando lo que los demás harían.
d) Observamos lo que otros hacen o han hecho para aprender una conducta social.

5. El fenómeno de la validación social, establece que (señala la incorrecta):

a) En los experimentos, este fenómeno depende del número de sujetos experimentales.
b) El número de individuos que actúan de una determinada forma va a determinar el número de personas que deciden imitarles.

c) Las campañas informativas pueden llegar a estimular aquello que pretenden evitar.

d) Las personas pueden verse inducidas a imitar la conducta del experimentador.

6. Un problema de las campañas informativas relacionado con la validación social es que:

a) Llaman la atención sobre un problema lamentando su excesiva frecuencia.

b) No llaman la atención sobre la forma de adecuarse a la norma establecida.

c) Se centran fundamentalmente en conductas negativas.

d) Implícitamente excluyen de la regla social las conductas que denuncian.

TEXT 2

La función del sueño en el aprendizaje

Aunque la función del sueño es desconocida, una hipótesis actual sostiene que podría intervenir en la consolidación de las huellas mnésicas en el cerebro adulto. Estas huellas mnésicas que se forman con el aprendizaje serían frágiles, fáciles de borrar o alterar, a menos que hubiera transcurrido un período de sueño. De hecho, muchas observaciones relacionan las fases de sueño con las huellas mnésicas.

Por ejemplo, los registros poligráficos muestran un aumento de la cantidad de sueño paradójico durante el sueño que sigue al aprendizaje, en el animal y en el hombre. A la inversa, la privación de sueño altera el aprendizaje de muchas tareas.

Pero la hipótesis que relaciona sueño y aprendizaje sigue siendo objeto de debate. Al fin y al cabo, podría ser que la consolidación de las huellas mnésicas del aprendizaje tuviera lugar durante la vigilia. Por ejemplo, se sabe que después de un aprendizaje espacial aparecen en el hipocampo de la rata unas ondas llamadas de “frente abrupto”. Durante las descargas de las ondas de “frente abrupto”, las poblaciones neuronales que estaban activas durante el aprendizaje se vuelven a activar, lo cual sugiere que las huellas mnésicas del aprendizaje son tratadas y eventualmente consolidadas. Ahora bien, estas ondas de frente abrupto no se producen sólo durante el sueño de ondas lentas, sino también en estado de vigilia tranquila.

El debate sobre la hipótesis que relaciona sueño y aprendizaje ha sido reactivado por observaciones de un estudio de percepción visual. En este estudio el sujeto tiene que detectar un elemento-diana presentado brevemente en un cuadrante de su campo visual. Si la prueba se repite el mismo día no se observa ninguna mejora de rendimiento. Después de una noche de sueño el sujeto mejora su ejecución y sigue siendo bueno al menos durante una semana, incluso a falta de todo nuevo entrenamiento.

Pero hay más. Si se priva de sueño al sujeto durante la primera noche, la que sigue al entrenamiento, no se produce ninguna mejora, ni siquiera si los sujetos duermen normalmente durante las noches siguientes. La primera noche parece pues obligatoria para consolidar la huella mnésica de aprendizaje que subyace al rendimiento del sujeto, al menos a nivel de la descripción comportamental.

QUESTIONS:

1. El texto se centra principalmente en discutir:

a) Dos hipótesis sobre el papel del sueño en la consolidación de huellas mnésicas.

b) La evidencia en favor de la hipótesis que relaciona sueño y aprendizaje.

c) Los mecanismos por los cuales el sueño facilita el aprendizaje.

d) El papel de las huellas mnésicas en el aprendizaje.

2. De los resultados discutidos en la primera parte del texto sobre el debate entre partidarios y detractores de la hipótesis del sueño, se puede concluir que el sueño:

a) Podría no tener una función obligatoria y exclusiva para la consolidación de las huellas mnésicas.

b) Es imprescindible para la consolidación de las huellas mnésicas.

c) Es necesario para la consolidación de las huellas mnésicas la primera noche después del aprendizaje, pero no las siguientes.

d) No parece ser imprescindible para la consolidación de las huellas mnésicas.

3. Para los defensores de la hipótesis que relaciona sueño y aprendizaje, el olvido se debería a:

a) Un fallo en la consolidación de la huella mnésica durante el sueño de ondas lentas.

b) Un fallo en la consolidación de la huella mnésica durante el aprendizaje.

c) Un fallo en la consolidación de la huella mnésica durante la vigilia tranquila.

d) Un fallo en la activación de la huella mnésica durante la fase de recuperación.

4. Una de las hipótesis sobre la función del sueño que se plantean en el texto afirma que:

a) El sueño podría intervenir en la consolidación de las huellas de memoria.

b) El proceso de aprendizaje se realiza fundamentalmente durante el sueño.

c) El proceso de aprendizaje interviene en la consolidación de las huellas mnésicas.

d) La fase del sueño de ondas rápidas participa en la consolidación del sueño.

5. Los registros poligráficos en el animal y en el hombre muestran:

a) Un aumento de la cantidad de sueño paradójico durante el sueño que sigue al aprendizaje.

b) Un aumento de la calidad de sueño durante el sueño post-aprendizaje.

c) Un aumento de las ondas lentas durante el aprendizaje.

d) Patrones de activación distintos en el animal y el hombre.

6. Durante la aparición de las ondas de “frente abrupto” en la rata, la huella mnésica:

a) Es consolidada, ya que las neuronas que estaban activas durante el aprendizaje se vuelven a activar.

b) Se consolida, ya que ha habido sueño post-aprendizaje.

c) Se borra a no ser que haya habido sueño pre-aprendizaje.

d) No se consolida durante la aparición de estas ondas durante la vigilia.

TEXT 3

Adicción a la televisión

Es pasmoso el número total de horas que la gente se pasa mirando la televisión. En el mundo industrializado la media de horas que la gente mira la televisión es de tres horas diarias, o sea, la mitad del tiempo libre. La cantidad de tiempo que la gente pasa ante el televisor preocupa a casi la mitad de la población, como demuestran sondeos de opinión.

Para explicar el preocupante número de horas que la gente pasa ante el televisor se han estudiado las reacciones a la televisión con experimentos de laboratorio. En ellos se registran las ondas cerebrales (mediante el electroencefalógrafo), la resistencia cutánea y el ritmo cardíaco de unos telespectadores.

En otro tipo de experimentos, para observar el comportamiento y las emociones de los telespectadores en el transcurso normal de la vida y no en las condiciones artificiales del laboratorio, los autores se sirvieron del "método de experiencias". En este método los participantes llevaron un avisador. Los experimentadores llamaban a los participantes al azar de seis a ocho veces diarias, durante una semana. Cada vez que los participantes oían la llamada anotaban en una tarjeta normalizada qué estaban haciendo y cómo se sentían en ese momento. Aquéllos que estaban viendo la televisión cuando los experimentadores les enviaban la señal contaban que se sentían relajados y pasivos. Era de esperar. Mientras se mira la televisión la estimulación mental es menor que cuando se está leyendo, como demuestran los estudios de laboratorio con electro-encefalogramas que miden la producción de ondas cerebrales alfa durante la realización de una tarea.

Lo más sorprendente es que la sensación de calma que describen los participantes de los experimentos se termina al apagar el televisor, pero no se elimina la pasividad. Que la televisión absorbe las energías es algo que los participantes suelen decir. Además, dicen que después de haber estado viendo la televisión tienen más dificultad para concentrarse que antes. En cambio, rara vez notan tal dificultad después de leer.

Estos efectos negativos de la televisión funcionan de un modo parecido a como lo hacen las drogas que crean hábito. Es mucho más probable que cause dependencia una droga que abandona el cuerpo enseguida que otra que permanezca en el cuerpo más tiempo. Con la droga que abandona el cuerpo enseguida la persona es más consciente de que sus efectos se van extinguiendo. De manera parecida, la vaga impresión que tienen los televidentes de que se sentirán menos relajados cuando dejen de mirar la televisión quizá sea una de las razones de que no dejen de verla. Ver televisión fomenta que se vea más televisión.

QUESTIONS:

1. Sobre la relación entre el consumo de drogas y la televisión, el autor defiende que ambos fenómenos**:**

a) Producen adicción por mecanismos psicológicos similares.

b) Comparten las vías neuronales implicadas en las adicciones.

c) No presentan relación respecto a la forma en que las personas consumen de forma adictiva.

d) Absorben la energía del adicto.

2. ¿Por qué crees que el autor afirma que "Ver televisión fomenta que se vea más televisión"?

a) Porque al dejar de ver la televisión la persona no se siente bien.

b) Porque las personas se sienten calmadas al ver la televisión.

c) Porque ver la televisión absorbe la energía del telespectador.

d) Porque las personas buscan ver la televisión para relajarse cuando están estresados.

3. De la opinión expresada por el autor se infiere que una buena terapia para combatir la adicción a la televisión podría ser:

a) Una terapia similar a la del consumo de drogas.

b) Una terapia centrada en el aumento de la actividad física.

c) Una terapia similar a la de la depresión.

d) Una terapia de relajación muscular mientras se ve la televisión.

4. Según el texto, la gente mira la televisión como media:

a) La mitad de su tiempo libre.

b) Un tercio de su tiempo libre.

c) Una cuarta parte de su tiempo libre.

d) Dos tercios de su tiempo libre.

5. Cuando se ve la televisión la producción de ondas cerebrales alfa:

a) Es menor que cuando se está leyendo.

b) Es mayor que cuando se está leyendo.

c) Es igual que cuando se está leyendo.

d) Está relacionada con un aumento en la estimulación mental.

6. La capacidad de concentración:

a) Disminuye después de ver la televisión.

b) Permanece igual después de ver la televisión.

c) Aumenta y / o permanece igual después de ver la televisión.

d) Es independiente de los períodos en los que se ve la televisión.

TEXT 4

¿Se acelera el cambio climático?

La continua sucesión de acontecimientos climáticos extremos, como huracanes, ciclones o tempestades, vuelve a poner de actualidad un problema ya recurrente en el debate sobre el cambio climático.

En el contexto del cambio climático, estos acontecimientos climáticos extremos tienen una resonancia especial y suscitan dos tipos de preguntas: en primer lugar, ¿estamos asistiendo a una aceleración del ritmo de este tipo de acontecimientos climáticos? Y, en segundo lugar, ¿esta aceleración se debe al impacto de las actividades humanas sobre el clima?

Apreciar la aceleración de los acontecimientos climáticos extremos asociados al cambio climático, y poder predecir los cambios en esta aceleración, requiere el análisis de series de datos tan largas y homogéneas como sea posible. Este análisis es complicado técnicamente debido a las transformaciones de las técnicas de medición con el paso del tiempo, al desplazamiento de las estaciones de observación e incluso a la progresiva densificación de la red de observación. En estas condiciones, es muy difícil establecer cualquier relación entre los acontecimientos climáticos extremos y las emisiones de gases por las actividades humanas.

Esto nos lleva al segundo aspecto del problema: ¿qué se sabe, en teoría, de las consecuencias de las actividades humanas en el ritmo de los acontecimientos climáticos extremos? ¿Hay que esperar una modificación de este ritmo? Para responder sobre la relación entre las actividades humanas y los acontecimientos climáticos extremos es necesaria la modelización numérica del clima.

En los modelos numéricos del clima más corrientes, la resolución gráfica es la mayoría de las veces del orden de algunos centenares de kilómetros. Este nivel de resolución permite contabilizar las depresiones, clasificarlas en función de su intensidad -las más profundas están asociadas a las tempestades- y detectar un cambio de tendencia en el ritmo de estos acontecimientos, pero no es suficiente para reproducir otro tipo de acontecimientos climáticos extremos como los ciclones tropicales y las tempestades de pequeñas dimensiones que pueden figurar entre las más intensas.

Un segundo método para modelar matemáticamente el clima consiste en simular la evolución de los parámetros estadísticamente relacionados con los acontecimientos climáticos extremos, que son de mayor escala y por lo tanto se pueden identificar en los modelos de baja resolución.

Por tanto, los principales obstáculos para el estudio de los acontecimientos climáticos extremos son la limitación en la resolución de los modelos matemáticos y la insuficiencia de las capacidades de cálculo, lo cual genera una descripción de los acontecimientos climáticos demasiado grosera. Además, las limitaciones de cálculo de los modelos obligan a hacer hipótesis estadísticas que no se sabe seguro si son válidas en las condiciones de un clima modificado por las emisiones de gases del cambio climático.

QUESTIONS:

1. Sobre la relación entre las actividades humanas y el cambio climático, el autor del texto defiende que la evidencia científica:

a) No permite conocer la relación entre las actividades humanas y el cambio climático, aunque futuros modelos podrían explicarla.

b) Demuestra una relación positiva entre ambas variables, aunque esta es limitada estadísticamente.

c) Está a favor de una relación estrecha entre ambos fenómenos existe, pero que estos datos son difíciles de analizar.

d) Muestra que el estudio de ambos fenómenos requiere modelos matemáticos por separado.

2. Sobre los modelos numéricos del clima, el autor sostiene que:

a) La deficiencia en la resolución gráfica de la tecnología actual limita su uso para el estudio del cambio climático, aunque puede ser útil para explicar ciertos fenómenos.

b) El descubrimiento de nuevos parámetros estadísticos puede mejorar el poder predictivo de la tecnología actual.

c) Son óptimos para la descripción y clasificación de fenómenos climáticos, pero limitados en su carácter predictivo.

d) Debido a la limitación en la resolución de la tecnología actual, son útiles para predecir fenómenos climáticos de carácter más intenso.

3. Si la tecnología para el estudio del cambio climático avanzara en la línea que demanda el autor, (señala la incorrecta):

a) Se podría reducir la aceleración del ritmo de los acontecimientos climáticos extremos.

b) Se podría determinar la relación entre los gases emitidos a la atmósfera y el cambio climático.

c) Se podría describir con precisión la evolución de ciclones y fenómenos similares.

d) Se podría comparar la evolución de los fenómenos climáticos en la actualidad con el clima en el pasado.

4. Los modelos matemáticos del clima permiten (señala la incorrecta):

a) Reproducir la conducta de ciclones tropicales.

b) Contabilizar determinados fenómenos atmosféricos.

c) Clasificar determinados fenómenos atmosféricos.

d) Detectar cambios de tendencia de las depresiones.

5. Según el texto, las hipótesis estadísticas derivadas de modelos del clima:

a) No se sabe seguro si son válidas en las condiciones de clima actuales.

b) Se han demostrado efectivas para la definición de los fenómenos problemáticos.

c) Se han demostrado erróneas por las limitaciones de resolución de los modelos.

d) Son fiables, pero requieren aún de validación empírica.

6. Los modelos matemáticos son limitados para:

a) La clasificación de tempestades de pequeñas dimensiones.

b) La clasificación de depresiones atmosféricas de pequeñas dimensiones.

c) La detección de cambios de tendencia de tempestades de grandes dimensiones.

d) La clasificación de depresiones atmosféricas de grandes dimensiones.

TEXT 5

Plantas resistentes: Arroz genéticamente manipulado

El taladro del arroz es un insecto cuya larva se introduce en el tallo de la planta del arroz y arrasa la producción de arroz. Para combatir el taladro del arroz se usan sustancias químicas, pero, como contrapartida, estas sustancias pueden ejercer un impacto negativo en los ecosistemas. Este impacto es más evidente en Europa, donde el arroz acostumbra a cultivarse en zonas de interés ecológico.

Una alternativa que puede contribuir de forma positiva al mantenimiento de la integridad de tales zonas de interés ecológico, además de ahorrar en costes de producción del arroz, consiste en utilizar variedades de arroz modificadas genéticamente mediante la introducción de genes de resistencia. Mediante esta transformación genética del arroz se han incorporado genes insecticidas del tipo *cry* (cry1B o cry1Aa) que codifican la expresión de las toxinas de Bacillus thuringiensis que combaten el taladro del arroz.

Este arroz modificado genéticamente se ha mostrado resistente al taladro del arroz, lo mismo en los ensayos de laboratorio e invernadero que en el campo. Además, se ha obtenido arroz modificado genéticamente con el gen inhibidor de proteasas del maíz (mpi). Este gen produce un significativo retraso en el crecimiento de las larvas del taladro del arroz, al bloquear su proceso digestivo.

El promotor de un gen controla cuándo y dónde se expresa dicho gen. Se busca, pues, que el gen se active sólo cuando se produzca el ataque del taladro del arroz y que el efecto del gen quede limitado al tejido afectado por el insecto. Así acontece con el promotor del gen mpi del maíz descrito anteriormente, que dirige la expresión de los genes insecticidas cry de suerte tal, que el efecto del gen sólo se produce en los tejidos atacados (heridos) por el taladro del arroz.

¿Qué decir del riesgo de flujo de genes entre variedades cultivadas y, sobre todo, desde éstas hacia el arroz salvaje? Aunque el arroz se multiplica por autofecundación, existe siempre cierta tasa de reproducción cruzada entre variedades de arroz cultivadas. Así pues, es necesario evaluar hasta qué punto existe el riesgo de que el polen de las variedades de arroz modificado genéticamente pueda fecundar plantas no modificadas, con la transmisión consiguiente de los genes incorporados.

QUESTIONS:

1. En tu opinión, cuál de las siguientes afirmaciones expresa mejor la idea principal del texto defendida por el autor:

a) Aunque la ingeniería genética es positiva para los alimentos, sus riesgos aún no se conocen suficientemente.

b) La ingeniería genética es beneficiosa para el cultivo del arroz y no tiene riesgos.

c) El autor sólo se centra en los beneficios de los cultivos transgénicos y ni siquiera nombra sus posibles riesgos.

d) El autor describe en detalle los pros y contras de la ingeniería genética de los alimentos.

2. De las palabras del autor sobre el gen *cry* se infiere que este gen:

a) afecta sólo a aquellas partes de la planta afectadas por el insecto.

b) funciona como una vacuna preventiva.

c) previene a aquellas partes críticas de la planta (tallo) de ser infectadas.

d) sólo puede introducirse en aquellas plantas de arroz infectadas.

3. En el texto se citan diversas variedades del arroz. Señala la afirmación correcta sobre las mismas:

a) El gen introducido en una variedad de arroz puede afectar a otra variedad distinta.

b) Los genes anti-taladro se encuentran en ciertas variedades del arroz.

c) Los genes anti-taladro sólo se reproducen en sus variedades de arroz originarias.

d) Los genes sólo pueden ser traspasados a otra variedad de forma artificial (ingeniería genética).

4. El taladro del arroz es un insecto:

a) Cuya larva se introduce en el tallo de la planta de arroz para atacarla.

b) Cuyo gen puede arrasar la producción de arroz.

c) Que al interactuar con ciertos genes pueden arrasar la producción de arroz.

d) Que se ha manipulado genéticamente para evitar que arrase la producción de arroz.

5. El promotor del gen *cry*:

a) Controla cuándo y dónde se expresa el gen.

b) Determina la capacidad reproductiva del gen.

c) Ataca exclusivamente al tallo de la planta de arroz.

d) Codifica la expresión de toxinas.

6. El gen descrito en el texto se activa cuando se produce el ataque y sólo:

a) en la zona del tejido atacada.

b) cuando el ataque es crítico para la planta.

c) en la zona de la planta donde reside el gen.

d) afecta al taladro del arroz, aunque se activa en la totalidad de la planta.

TEXT 6

Plantas resistentes: Arroz genéticamente manipulado.

El taladro del arroz es un insecto cuya larva se introduce en el tallo de la planta del arroz y arrasa la producción de arroz. Para combatir el taladro del arroz se usan sustancias químicas pero, como contrapartida, estas sustancias pueden ejercer un impacto negativo en los ecosistemas. Este impacto es más evidente en Europa, donde el arroz acostumbra a cultivarse en zonas de interés ecológico.

Una alternativa que puede contribuir de forma positiva al mantenimiento de la integridad de tales zonas de interés ecológico, además de ahorrar en costes de producción del arroz, consiste en utilizar variedades de arroz modificadas genéticamente mediante la introducción de genes de resistencia. Mediante esta transformación genética del arroz se han incorporado genes insecticidas del tipo *cry* (cry1B o cry1Aa) que codifican la expresión de las toxinas de Bacillus thuringiensis que combaten el taladro del arroz.

Este arroz modificado genéticamente se ha mostrado resistente al taladro del arroz, lo mismo en los ensayos de laboratorio e invernadero que en el campo. Además, se ha obtenido arroz modificado genéticamente con el gen inhibidor de proteasas del maíz (mpi). Este gen produce un significativo retraso en el crecimiento de las larvas del taladro del arroz, al bloquear su proceso digestivo.

El promotor de un gen controla cuándo y dónde se expresa dicho gen. Se busca, pues, que el gen se active sólo cuando se produzca el ataque del taladro del arroz y que el efecto del gen quede limitado al tejido afectado por el insecto. Así acontece con el promotor del gen mpi del maíz descrito anteriormente, que dirige la expresión de los genes insecticidas cry de suerte tal, que el efecto del gen sólo se produce en los tejidos atacados (heridos) por el taladro del arroz.

¿Qué decir del riesgo de flujo de genes entre variedades cultivadas y, sobre todo, desde éstas hacia el arroz salvaje? Aunque el arroz se multiplica por autofecundación, existe siempre cierta tasa de reproducción cruzada entre variedades de arroz cultivadas. Así pues, es necesario evaluar hasta qué punto existe el riesgo de que el polen de las variedades de arroz modificado genéticamente pueda fecundar plantas no modificadas, con la transmisión consiguiente de los genes incorporados.

QUESTIONS:

1. La idea principal del texto consiste en:

a) La propuesta de una explicación a las propiedades anómalas del agua.
b) La descripción de propiedades anómalas del agua.
c) La descripción de las propiedades químicas del agua.
d) La propuesta de una explicación para las propiedades microscópicas del agua.

2. ¿Cuál de las siguientes afirmaciones podría corresponder a la opinión del autor del texto?:

a) La importancia del agua no se ha visto correspondida con suficiente investigación.

b) Las propiedades del agua han sido estudiadas ampliamente.

c) La importancia del agua no se corresponde con la simplicidad de sus propiedades.

d) El estudio de las propiedades del agua permite entender sus mecanismos en profundidad.

3. Una vez que el agua se convierte en un estado particular (por ejemplo, hielo):

a) Su estructura varía incluso sin cambiar de estado.

b) Su estructura de hidrógeno no varía con el tiempo.

c) Su estructura de hidrógeno varía sólo si pasa a otro estado (por ejemplo, de líquida a hielo).

d) Su estructura puede variar incluso si su energía permanece constante.

4. El agua en su estado sólido (el hielo):

a) Es menos densa que la líquida.

b) Es menos densa que la fase metaestable y supercongelada.

c) Es menos densa que la fase supercalentada.

e) Es más densa que la fase líquida y supercalentada.

5. El entramado de enlaces de hidrógeno y su dinámica:

a) Determinan las propiedades del agua.

b) Interacciona con las propiedades del agua.

c) Interacciona con la densidad del agua para determinar su fase.

d) Determina la fase en que se encuentra el agua.

6. En relación con las propiedades del agua:

a) No existe una explicación única para la mayoría de las mismas.

b) Existen fenómenos poco estudiados.

c) Existen hipótesis controvertidas para su explicación.

d) Se conocen en su mayoría debido a la simplicidad de su estructura.

B2. Questionnaire on medium preference:

Estimado o estimada participante:
Nos gustaría que cumplimentases este formulario. Consta en primer lugar de tres preguntas abiertas.
Posteriormente deberás valorar una serie de afirmaciones marcando un valor del 1 al 5, donde:
1 = completamente en desacuerdo; **5 = completamente de acuerdo**.
MUCHAS GRACIAS.

ID: __________________ Edad: _______ Ocupación: ___________________

a. ¿A qué edad comenzaste a usar el ordenador u otros dispositivos como tablet o smartphone? ______

b. Aproximadamente, ¿cuántas horas al día usas el ordenador, la tablet o el smartphone? ________

c. De esas horas, ¿cuánto tiempo usas estos dispositivos para trabajar o estudiar? __________

	En desacuerdo				**De acuerdo**
	1	2	3	4	5
1. Prefiero estudiar para mis asignaturas con el libro de texto o el manual en papel antes que en formato electrónico.	O	O	O	O	O
2. Cuando encuentro en Internet un texto que necesito para realizar un trabajo para alguna asignatura lo leo en el ordenador, la tablet o el smartphone en lugar de imprimirlo.	1 O	2 O	3 O	4 O	5 O
3. Entiendo y memorizo mejor cuando estudio leyendo en formato electrónico que cuando leo en papel.	1 O	2 O	3 O	4 O	5 O
4. Cuando tengo que reflexionar sobre varios textos de un tema que estoy estudiando y organizar la información, prefiero imprimirlos antes que leerlos en algún dispositivo electrónico.	1 O	2 O	3 O	4 O	5 O
5. Si tengo apuntes de alguna asignatura en un archivo escritos a ordenador, para mí es mejor estudiarlos en algún dispositivo electrónico o el ordenador que imprimirlos para estudiarlos en papel.	1 O	2 O	3 O	4 O	5 O

Appendix C

Materials used in Study 3

C1. The text used in this study is available at

https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/robots-que-aprenden-como-nios-735/robots-autodidactas-16312

C2. Comprehension questions[22]

1. ¿Cuál de las siguientes opciones representa mejor la idea principal del texto?:

   a. Entender cómo aprende el cerebro humano nos permitirá desarrollar robots que sean tan inteligentes como las personas.

   b. Desarrollar robots que aprenden de forma autónoma ayudará a nuestra comprensión sobre cómo aprende el cerebro humano.

   c. Para comprender la realidad los robots deberán crear predicciones que minimicen su discrepancia respecto a dicha realidad.

   d. Aumentar nuestro conocimiento sobre inteligencia artificial nos permitirá crear robots que puedan aprender de forma autónoma.

2. Según el modelo de aprendizaje expuesto en el texto, ¿qué transporta la señal que retorna a los niveles de procesamiento superior del cerebro?

   a. Las predicciones acerca de la realidad

   b. Los conocimientos previos sobre la realidad.

   c. La discrepancia entre lo estimado y lo percibido.

   d. La información percibida por los sentidos.

[22] Questions 2, 3, 5, 8, 11, and 21 were excluded in order to increase the reliability of the test.

3. La capacidad de un sistema robótico denominada "motivación intrínseca" hace que busque de forma autónoma las situaciones con mayor potencial de aprendizaje. Esta determinación posibilita que aprenda:

a. De manera azarosa al interactuar con el entorno.

b. Aquello que le motiva según su programación.

c. A través de recompensas o refuerzos externos.

d. A no cometer predicciones erróneas al interactuar.

4. Al respecto de la conducta social, la investigación con inteligencia artificial sugiere que los humanos realizamos algunas conductas elementales de ayuda a los demás porque:

a. Aprendemos las normas sociales que predicen qué es lo correcto.

b. Tratamos de conseguir que nuestras predicciones se confirmen.

c. Reciben recompensas sociales al predecir la conducta de los demás.

d. Hemos aprendido a predecir las normas sobre el mundo social.

5. Para que las oportunidades de aprendizaje sean más altas, los niños tienden a elegir situaciones de aprendizaje en las que:

a. Pueden realizar hipótesis fáciles de verificar.

b. Se topan con algo desconocido.

c. Encuentran mucha riqueza de estímulos.

d. Cometen pocos errores de predicción.

6. Los sistemas de inteligencia artificial a los que se les incorpora un cuerpo pueden asociar acciones físicas a la información que reciben. Se ha comprobado que esto facilita que dichos sistemas:

a. Desarrollen capacidades sensoriales.

b. Creen algoritmos que determinen dichas acciones.

c. Reconozcan las características de una imagen.

d. Desarrollen mejores habilidades numéricas.

7. Los estudios sobre inteligencia artificial podrían ser beneficiosos en relación con nuestro conocimiento de los trastornos del desarrollo. Al respecto, el texto sugiere una hipótesis que propone que las personas con autismo prefieren las rutinas repetitivas porque:

a. Así aprenden a minimizar los errores en sus predicciones.

b. No son capaces de detectar errores en sus predicciones.

c. No tienen habilidades cognitivas para realizar predicciones.

d. Son demasiado sensibles a los errores de sus predicciones.

8. Los robots cachorros de Sony son programados de forma que buscan por sí mismos distintas situaciones de aprendizaje, lo que afecta a su nivel de exploración del entorno. Esto sugiere que el tipo de habilidades que alcanzan las personas depende:

a. Tanto de la carga genética como del entorno.

b. Del tipo de interacciones que realiza con el entorno.

c. De su habilidad para reducir los errores de predicción.

d. De la cantidad de predicciones acertadas.

9. Si se produce falta de concordancia entre las señales de los niveles superiores del cerebro y las de los sentidos, el sistema de procesamiento de la información detecta errores de estimación que:

a. Dificultan la adquisición de conocimiento.

b. No permiten conocer realmente la realidad.

c. Mejoran nuestros modelos acerca de la realidad.

d. Procesan los rasgos básicos de la realidad.

10. Según el modelo de aprendizaje propuesto en el texto, disponemos de un cerebro predictivo, puesto que casi todas nuestras acciones se basan en la generación y actualización de predicciones. Este proceso se dirige a:

a. Minimizar el número de predicciones generadas.

b. Comprobar todas las predicciones generadas.

c. Mejorar el ajuste entre las creencias previas y lo percibido.

d. Percibir la realidad para generar predicciones futuras.

11. Se ha propuesto desarrollar robots cuyo sistema de inteligencia artificial permitiría comprender mejor el trastorno por déficit de atención e hiperactividad. Para ello, dicho sistema estará diseñado para que su procesamiento predictivo:

a. Imite a los humanos tras aprender interactuando con ellos.

b. Tenga preferencia por la estimulación impredecible.

c. Se comporte según lo que sabemos de dicho trastorno.

d. No sea capaz de mantener su atención cuando interactúa.

12. En el texto se afirma que nuestro cerebro trata constantemente de predecir el futuro. Esto es posible debido a que el cerebro:

a. Produce estrés ante la incertidumbre de lo desconocido.

b. A veces no es capaz de determinar qué es aquello que percibimos.

c. Refina sus modelos de la realidad según la información de los sentidos.

d. Interpreta objetivamente la información de la realidad

13. Se propone que para acercar los sistemas de inteligencia artificial a la inteligencia humana sería necesario que su proceso de aprendizaje y de desarrollo se produzca como un complejo “sistema de cascadas”. Con ello se quiere representar que es necesario que la inteligencia artificial adquiera conocimiento complejo:

a. Sumando conocimientos más simples.

b. Minimizando los errores de cada predicción.

c. Paso a paso, sobre la base de lo aprendido antes.

d. Realizando correctamente todas las predicciones.

14. Tanto para la psicología del desarrollo como para la robótica, un área clave en la investigación sobre el aprendizaje es conocer cómo los humanos aprendemos a andar, hablar y a reconocer objetos y personas. Su estudio es muy relevante porque son habilidades:

a. Que se aprenden de manera autónoma.

b. Fundamentales de nuestra conducta.

c. Que nos diferencian de los animales.

d. Fundamentales para ayudar a los demás.

15. En las reflexiones presentadas en el texto acerca de los robots del futuro, se concluye que los robots capaces de desarrollar una inteligencia similar a la humana:

a. Se construirán cuando conozcamos bien cómo funciona el cerebro.

b. Son ya una realidad en el campo de la investigación robótica.

c. Se han conseguido ya en juegos como el ajedrez o el go.

d. Deben considerarse como una realidad aún bastante lejana.

16. Algunos investigadores indican que el cuerpo de un robot modifica decisivamente el proceso de aprendizaje del sistema de inteligencia artificial que lo controla. Esto se debe a que:

a. Su aprendizaje está determinado por sus interacciones corporales.

b. Sus algoritmos deben determinar los movimientos de su cuerpo.

c. Su aprendizaje depende de un proceso más lento.

d. Puede ser programado de manera distinta en cada experimento.

17. En el texto se menciona que ya existen sistemas de inteligencia artificial que superan las capacidades de los humanos para jugar al ajedrez o al juego chino *go*. Al respecto en el texto se concluye que este hecho:

a. Demuestra que las máquinas pueden ser más inteligentes que los humanos.

b. No es comparable a la capacidad humana para el aprendizaje espontáneo.

c. Indica que los humanos cometemos errores al tratar de predecir la realidad.

d. Sugiere que en 10 o 20 años las máquinas serán tan inteligentes como los humanos.

18. Los robotistas intentan crear máquinas que imiten el desarrollo espontáneo de los niños. ¿Qué aporta esto en su colaboración con psicólogos y neurocientíficos para el estudio del desarrollo humano?

a. Crean robots aptos para aplicar terapia a personas con autismo.

b. Descubren nuevos factores que ayudan a entender cómo aprendemos.

c. Identifican los efectos de la genética en el desarrollo humano.

d. Encuentran formas para intentar superar a la inteligencia humana.

19. Los estudios encuentran que los niños con bajas habilidades numéricas y de cálculo tienen dificultades para:

a. Calcular con los ojos cerrados.

b. Crear imágenes mentales de sus dedos.

c. Asociar los dedos a cantidades.

d. Calcular sin usar los dedos.

20. En unos experimentos en los que trataban de enseñarle vocabulario a robots, descubrieron que aprendían más fácilmente una palabra si al nombrar el objeto correspondiente este se colocaba:

a. Junto a otros objetos muy distintos.

b. Donde el robot pueda moverse para mirarlo.

c. Donde indique el algoritmo de su sistema.

d. En la misma posición cada vez.

21. Según la explicación propuesta en el texto para el trastorno por déficit de atención e hiperactividad, las personas con dicho trastorno tendrían problemas para concentrarse porque:

a. No pueden mejorar sus predicciones de la realidad.

b. Perciben la información de la realidad con demasiada rapidez.

c. Tienen el sistema nervioso excesivamente activado.

d. Sienten una atracción constante por la estimulación impredecible.

C3. Questionnaire on previous knowledge and topic interest

ID PARTICIPANTE: ________

Indica del 1 al 10 cuánto crees que sabes de los siguientes temas:
(1 = Nada; 10 = Conocimiento experto)

1. Las nuevas tecnologías
2. Investigación y desarrollo de robots:
3. Investigación y desarrollo de inteligencia artificial:
4. Programación informática:
5. Aprendizaje infantil:
6. Psicología del aprendizaje:
7. Psicología del desarrollo:
8. Procesos cerebrales implicados en el aprendizaje:

Indica del 1 al 10 cuánto te interesan los temas relacionados con:
(1 = Nada; 10 = Lo máximo que puede interesarte algo)

1. Nuevas tecnologías:
2. Investigación y desarrollo de robots:
3. Investigación y desarrollo de inteligencia artificial:
4. Programación informática:
5. Aprendizaje infantil:
6. Psicología del aprendizaje:
7. Psicología del desarrollo:
8. Procesos cerebrales implicados en el aprendizaje:

[This page intentionally left blank]

"*Know what's weird? Day by day nothing seems to change, but pretty soon, everything is different*."

Calvin tells Hobbes, his best friend.

(by Bill Watterson)