

To cite this manuscript:

Altamura, L., Vargas, C., & Salmerón, L. (2023). Do New Forms of Reading Pay Off? A Meta-Analysis on the Relationship Between Leisure Digital Reading Habits and Text Comprehension. *Review of Educational Research*, 0(0). <https://doi.org/10.3102/00346543231216463>

Do new forms of reading pay off? A meta-analysis on the relationship between leisure digital reading habits and text comprehension

Lidia Altamura
Cristina Vargas
Ladislao Salmerón

University of Valencia

Abstract

Previous research has evidenced a strong positive relationship between leisure print reading habits and reading comprehension across the lifespan. The rapid evolution of new forms of leisure digital reading could modify such relationship. This meta-analysis extends previous research by analyzing the relationship between leisure digital reading habits and reading comprehension. We analyzed 40 effect sizes using multilevel analysis. Data involved 469,564 participants from studies published between 2000-2022. The average effect size reflects a small significant effect on reading comprehension ($r = .055$), which contrasts with the medium size effects found in the literature related to print reading habits and comprehension. This relationship is significantly moderated by the readers' educational stage. At early stages (primary and middle school) negative relationships are observed between leisure digital reading and text comprehension, while at later stages (high school and university) the relationship turns positive. We highlight the different contributions that reading modalities and technological contexts have to our reading comprehension, especially across the lifespan. In sum, leisure digital reading does not seem to pay off in terms of reading comprehension, at least, as much as traditional print reading does.

Keywords: digital reading, reading habits, text comprehension, meta-analysis

"Educational impact and implications statement."

There are positive associations between the reading frequency of printed texts and comprehension. We found that such association is almost null when it comes to reading digital texts.

Author Note: Funding for this research was provided by Grant PID2020-118512GB-I00 funded by MCIN/AEI/10.13039/501100011033, financed by the Spanish Ministry of Science and Innovation. Correspondence concerning this article should be addressed to Ladislao Salmerón, University of Valencia. Contact: Ladislao.Salmeron@valencia.edu

Decades of research have evidenced solid benefits of reading habits in terms of reading comprehension development across the lifespan (Locher & Pfof, 2020; Mol & Bus, 2011; van Bergen et al., 2020). Evidence converges on the positive reciprocal relationship between reading practices and the development of reading comprehension skills (Stanovich, 1986). As young readers are exposed to increasingly longer and more complex texts, they further develop their ability to comprehend texts, which in turn, encourages them to pursue even more complex texts. Such relationship, that has been established for habits involving printed texts, may be different for new forms of reading. Currently, reading not only takes place on paper. The growing evolution of digital reading devices, the constant access to such devices, and the new types of reading materials available (e-books, social networks, digital newspapers...) have suddenly introduced new reading habits. In terms of leisure time, reading frequency for print materials has decreased in the last decade. A comparison of PISA data from 2012 and 2018 showed that adolescents voluntarily read fewer fiction books, magazines, or newspapers. On the contrary, they read more to fulfill their practical needs, and they read more using online formats, such as chats, online news or websites containing practical information (e.g., schedules, events, tips, recipes) (OECD, 2019). Studies suggest that the increasing access to digital devices is taking place at the expense of traditional reading of printed books (Merga & Roni, 2017). A national study examined data from 1,021,209 students of eighth, tenth, and 12th grades in the United States, from 1976 to 2016 (Twenge et al., 2018). The authors support the displacement theory after evidencing a substantial decline in book leisure reading and a sharp growth of digital media consumption across all ages. As online texts tend to have rather low linguistic qualities (Snow, 2010) and are often created for quick and brief consumption, such increase in digital reading frequency could impact readers' development of comprehension skills. Wolf (2018) has recently hypothesized that such influence should be particularly evident during early years of Primary school, where the focus of reading instruction is still on developing reading foundations (decoding, fluency or vocabulary) and not on reading to learn.

The emergent popularity of these new forms of reading has brought a renewed emphasis on the study of the relationship between reading habits and reading comprehension, but, to the best of our knowledge, no previous study has synthesized the existing literature. Therefore, we present a meta-analysis of recent studies that have specifically investigated the relationship between digital reading habits and reading

comprehension, focusing on leisure reading (not motivated by academic or work-related purposes). In addition, we explore the effects of several potential moderator variables.

Print reading habits

Print reading habits usually involve reading novels, magazines, newspapers, or comics on paper. Such habits have been positively associated with reading comprehension development. In their meta-analysis, Mol and Bus (2011) studied the relationship between reading habits, measured by print exposure level, and reading comprehension skills from infancy to early adulthood. From the 99 studies analyzed, moderate to strong correlations were found (r for Grades 1 to 12 = .36), being the highest in higher-education readers (r = .41).

Potential underlying mechanisms for this relationship consist of a “virtuous circle of reading” in which time spent reading and reading comprehension are self-reinforced. One well-established model to explain such relationship is the Matthew effect by which “the rich get richer, and the poor get poorer” statement is applied to reading development (Pfof et al., 2014; Stanovich, 1986). In this way, children who are good readers tend to engage in more frequent and more challenging reading activities, thus enhancing their reading comprehension at a faster pace than poor readers, who usually do not find much pleasure in reading and thus improve their skills at a slower pace. Recent longitudinal studies have analyzed the influence of print reading habits on the development of reading skills throughout childhood (Torppa et al., 2019; van Bergen et al., 2020) and from pre-adolescence to later adulthood (Locher & Pfof, 2020), providing empirical support for a Matthew effect on reading.

Digital reading habits

As research on digital reading habits is still emerging, definitions of what constitutes such habits widely vary across studies. Definitions tend to emphasize purposes, text types and reading frequency. Regarding purposes, authors tend to distinguish between academic or school and leisure reading habits (Torppa et al., 2019), also referred to as extracurricular (Pfof et al., 2014) or recreational (McGeown et al., 2016). Regarding texts, major types of digital texts distinguish between on-screen texts (non-interactive texts presented on digital devices such as tablets, or computers); and hypertexts and hypermedia (interactive and linked texts, which may or not be connected to the Internet) (Coiro, 2021). Finally, digital reading habits are usually measured in

terms of temporal frequency via self-reports (e.g., “How often do you read...?”). Accordingly, we define leisure digital reading habits as the time spent on digital texts for the purposes of social-communication (e.g., instant messaging, online chatting, emailing, checking social media) or informative-linear reading (e.g., looking for information on the Internet, browsing a website, blog or forum, reading e-books, e-magazines or e-comics) (Duncan et al., 2016; McGeown et al., 2016).

The emerging availability of digital reading devices during the 2000s, and the subsequent increase in leisure digital reading habits in the population, have raised concerns about the well-established relationship between reading habits and comprehension. Within the existing literature, we find evidence pointing in two opposite directions. On the one hand, longitudinal studies evidence a negative relationship between digital text reading habits (measured as reading frequency via self-questionnaires and parental reports in Grades 1 to 4) and reading comprehension skills at later ages (from Grade 6 to 7) (Torppa et al., 2019). In a bidirectional way, lower reading comprehension predicts higher frequency of digital reading in earlier stages (from Grade 4 to 6). Cross-sectional studies also report negative relationships between digital reading habits and reading comprehension regardless of the purpose of reading: leisure activities (Naumann & Sälzer, 2017; Park & Weng, 2020) or academic tasks (Amiama-Espaillat & Mayor-Ruiz, 2017; Gumus & Atalmis, 2011; Park & Weng, 2020). Recent meta-analyses can help understand the mechanisms that are involved in these negative relationships, as they converge on the existence of a print-reading superiority effect. This phenomenon states that worse text comprehension is observed when reading on a digital device compared to reading the same text in print. This effect has been evidenced in older readers, from post-secondary students to undergraduates, (Clinton, 2019; Delgado et al., 2018; Kong et al., 2018), and in pre-readers (Furenes et al., 2021). Even though the effect sizes are small (*Hedges’ g* between -.21 and -.25), one could argue that in the midterm, the detrimental effect of continuously reading on digital devices could prevent young students from fully developing their potential reading comprehension (Salmerón et al., 2023).

Two unique factors suggest that spending leisure time with digital reading tools may not pay off in terms of reading comprehension development. First, digital reading devices can serve many different purposes other than reading, which incentivizes multitasking and distractions from reading (Baron et al., 2017), with potentially detrimental consequences for comprehension (Clinton-Lisell, 2021; Liu & Gu, 2020;

Nikkelen et al., 2014). Second, the Internet has brought new types of reading with particular features such as short and fast-paced stimuli, low-quality content and informal language. According to the *shallow hypothesis* (Annisette & Lafreniere, 2017; Carr, 2010) the frequent use of ultra-brief social media promotes rapid and relatively shallow thought, associated with cognitive and moral shallowness. When it comes to reading, superficial processing, such as scanning and skimming (Mangen, 2008), refers to the gradual shift that digital readers are experiencing towards higher reading frequency of short-length texts and fast-paced content, which do not tend to foster engagement nor deep processing (Baron, 2021; Ra et al., 2018). In addition, digital readers are highly exposed to non-professionally created content written by laypersons, with the subsequent risk of encountering fragmented, decontextualized and even inaccurate information (Bråten et al., 2018). Accordingly, one could argue that exposure to such low-quality text material may not prepare readers to fully engage in deep comprehension processes. Lastly, the type of language that is usually used in digital texts may not follow vocabulary and grammar rules. Contrary to the well-established positive association between print reading and reading comprehension, characteristics of the vocabulary presented in digital texts, which are closer to the oral language than the written one (Snow, 2010), could not display such positive relationship. For instance, social media, online messaging, or blogs often have conversational traits and are written in informal language. Therefore, digital text readers may not be exposed to the academic vocabulary and high-level grammar structures found in traditional printed texts.

On the other hand, there is evidence suggesting a potential positive relationship between digital reading habits and comprehension. Some studies evidenced positive associations between digital reading frequency and reading comprehension in adolescents (Botello-Peñaloza & López-Alba, 2014; Duncan et al., 2016; Islam et al., 2020; Jackson et al., 2011; Johnson, 2013; Kong et al., 2022; Wu & Peng, 2017) as well as in college students (Amiama-Espaillet & Mayor-Ruiz, 2017; Chen & Fang, 2013; Mežek, 2013). This could be partially explained because digital reading presents exclusive features that we cannot encounter in print materials such as interactivity, nonlinearity, immediacy in accessing information and the convergence of text, images, audio, and video (Chen & Chen, 2014). Interactive features of digital devices, such as questions with feedback, digital glossaries, and collaborative tools could benefit reading comprehension performance, as a recent meta-analysis developed by Clinton-Lisell et

al. (2021) has shown ($g = .66$). Moreover, readers usually find the opportunity to search inside and across different materials beneficial (Baron et al., 2017). Digital reading is often more practical, which is why readers tend to access digital materials to pursue a goal (e.g., gaining more information about a topic, checking for timetables, recipes...) (Stadtler et al., 2013), which can potentially increase reading motivation and engagement (Assor et al., 2002). Furthermore, online reading also encompasses communication and social functions (e-mails, instant messaging, social media) and allows social and collaborative reading (Leu et al., 2011).

Previous literature has analyzed reading comprehension and reading habits across different populations, for example, participants of different ages or educational stages. Looking at specific individual differences could help to clarify the existing differences in the literature on the relationship between digital reading habits and reading comprehension.

Individual characteristics and the development of reading comprehension

Readers' age is one of the most salient factors in the literature analyzing the relationship between reading habits and text comprehension. Most studies have focused on print reading habits. Through several longitudinal studies, it has been evidenced that such relationship is not constant and may change across the lifespan. During secondary school, the relationship between reading habits and reading comprehension becomes more relevant, but it weakens as students grow older, reaching a stable level in adulthood. The longitudinal study by Locher and Pfof (2020) showed a medium-sized correlation between leisure reading habits and reading comprehension for students attending secondary school ($\beta = 0.20$; average age = 14.67), i.e., a sample from fifth to ninth grade. The correlations decreased with age ($\beta = -0.04$, in the college cohort; average age = 24.77) and reached a stable low level in adulthood ($\beta = 0.07$; average age = 50.74). In a different longitudinal study, van Bergen et al. (2020) analyzed data from 200 Finnish children, followed from age 5 to age 15. Path analyses showed that children's print reading habits and reading comprehension were bidirectionally related. During the early primary school years (second and third grade), the effects run from reading comprehension to print reading habits ($r = .28$), but not the other way around ($r = .12$, not significant). The effect of accumulated -but not recent- practice only emerged in adolescence ($r = .38$). Indeed, this significant relation comes from Grade 3 to 9, but not from Grade 7 to 9. Torppa et al. (2019) reported similar findings from a longitudinal

study that followed a sample of 2,525 Finnish students, from age 7 to 16. In Grades 1 to 3, better reading comprehension and fluency predicted more frequent leisure reading, while in later Grades (6 to 9) more leisure reading, particularly of books, predicted higher reading comprehension growth. When it comes to digital reading habits, recent cross-sectional studies emphasize the differences between age cohorts. Salmerón et al. (2023), analyzed NAEP data from 2017 (National Assessment of Educational Progress, n.d.), and found stronger negative relationships between the frequency of use of digital devices in the classroom and reading comprehension in fourth grade, compared to eighth grade (unstandardized estimates = -4.03 vs -2.07).

In sum, there is no clear pattern for the relationship between leisure digital reading habits and the development of reading comprehension. Apart from individual differences, the readers' social and cultural context could also shape the effects of reading habits on reading comprehension, as we review next.

Technological context and digital reader's habits

Recent years have seen enormous technological advancements, and with these changes has come a transformation in people's reading habits. Social media, a key element in new reading forms, was gradually introduced during the mid-2000s. Similarly, tablets were only available from 2010 on. With these changes, reading habits have also undergone changes, transitioning from more traditional formats (print) to more modern ones (digital devices). In principle, as long as readers continue reading, one could expect no major change in the size of the relationship between reading habits and comprehension across years. According to the displacement hypothesis (Neuman, 1988), this is because time previously devoted to print reading may just be replaced by time devoted to digital reading.

On the contrary, from the perspective of the *shallowing hypothesis* (Annisette & Lafreniere, 2017; Carr, 2010), we may expect that the relationship between digital reading habits and comprehension diminishes across years. As social media consumption grows, people have less time for more productive forms of digital reading. Consequently, the time spent reading digital texts may not be as beneficial as it was in previous years. Empirical evidence for this assumption comes from the meta-analysis from Delgado et al. (2018) regarding the differences between digital and print reading on text comprehension. Authors regressed the effect size of that comparison, coded in a way so that positive values indicated a higher difference in favor of print, on the year of

publication of the original studies, ranging from 2000 to 2017. Results revealed a small significant positive effect, indicating that over the years the difference in favor of print was growing, but not diminishing, as one could have expected if the effect was due to lack of practice with digital tools in early studies. Authors interpreted this pattern as evidence for a growing negative effect of current digital reading practices, highly associated with quick interactions with short texts. Therefore, the moment in which a study takes place allows us to locate the individual in a given time period, providing us with information about the current social situation, and helps us identify potential generational characteristics and differences.

The current study

To the best of our knowledge, our study is the first meta-analysis to analyze the relationship between leisure digital reading habits and reading comprehension. Related meta-analyses, already discussed above, have explored the relation between print reading habits (Mol & Bus, 2011) or reading media (Clinton, 2019; Delgado et al., 2018; Furenes et al., 2021; Kong et al., 2018) on comprehension.

The inconsistency across existing literature when it comes to the relation between digital reading habits and comprehension has led us to carry out this meta-analytic study. With this explorative meta-analysis, we aim to determine the relationship between digital reading habits and reading comprehension, evidenced by empirical research conducted during the last two decades. The research objectives of this work are:

- 1) To determine the most prevalent association between leisure digital reading habits and reading comprehension. According to the *displacement hypothesis*, we could expect a positive association between digital reading habits and reading comprehension, as students may just be replacing the time previously devoted to print reading to digital reading. On the contrary, based on the *shallowing hypothesis* we could expect negative to null effect sizes, due to the frequent low linguistic quality, short-length and fast-paced content of digital texts.
- 2) To identify potential moderator variables for the relationship between leisure digital reading habits and reading comprehension, including age, year of data collection, and type of digital reading habits. Based on the reviewed literature of individual characteristics and the development of reading comprehension, we

were unable to specify a clear pattern. Accordingly, we test the effect of age and educational stage as an exploratory research goal. In addition, we hypothesize that more recent studies will present a more negative association between leisure digital reading habits and reading comprehension. Lastly, we expect that the relationship between digital reading habits and reading comprehension will be stronger for linear-informative digital reading habits, compared to social-communicative ones.

Method

In this section, the search procedure, selection, and codification of the included studies are described.

Search

We conducted an initial systematic search using several techniques. The first one corresponded to looking through some academic databases: Web of Science, Scopus, Psych Info and Eric. Web of Science, Psych Info, and Eric databases also contain gray literature such as dissertations. We entered the following search equation: (“Reading frequency” OR “leisure reading” OR “reading activity” OR “Reading habits” OR “digital habits” OR “print reading” OR “print exposure” OR “social media use” OR “ICT use” OR “internet use”) AND (“text comprehension” OR “Reading comprehension” OR “Reading capacity” OR “reading skills” OR “Reading abilities” OR “literacy level”). The publication years were restricted from 2000 to 2022. These terms were searched as title, abstract, or keywords. We added a second specific search for studies that have worked with PISA or PIRLS databases with the search equation: (“PISA” OR “PIRLS”) AND (“computer” OR “digital” OR “online”) AND (“reading”).

As recommended by Card (2012), we complemented the search with additional strategies. Therefore, in a second phase, we examined references included in previous reviews (3 eligible studies were found in these reviews but, they were already identified in the initial search). Next, we approached experts and societies in this area (The Society for Text and Discourse, European Association for Learning and Instruction, American Educational Research Association) asking for information about unpublished studies. Additionally, a forward search was performed using Google Scholar to find studies that cited the selected works. Finally, references from the selected studies were also retrieved. The search ended in January 2022.

To filter the search results, we selected those studies that met the following inclusion criteria:

1. It is a scientific study in which digital reading habits and reading comprehension are measured with no intervention (e.g., As an example of excluded work, in the study of Mak et al., 2019, the aim was to analyze the effect of a gamified tool in leisure reading habits).
2. Reading comprehension is evaluated in an objective way, with standardized (e.g., Petko et al., 2017 who used PISA 2012 data) or ad-hoc tests, i.e., those created by the researchers for their study purposes. As an example of an excluded work, in the study of Bann et al., 2006 participants self-reported their reading comprehension level.
3. The reading habits assessed referred to leisure time and required reading from some digital device, not only watching (TV) or playing on them (video games). As an example of an excluded work, Jackson et al. (2011) analyzed the relations between general Internet use, video game playing and cell phone use on reading skills.
4. The study makes an empirical contribution that includes the results of the comparison (i.e., the paper is not a review or an opinion).
5. The study was published or presented between 2000 and 2022. Formal publication was not required.
6. The report was written in English or Spanish.
7. 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).

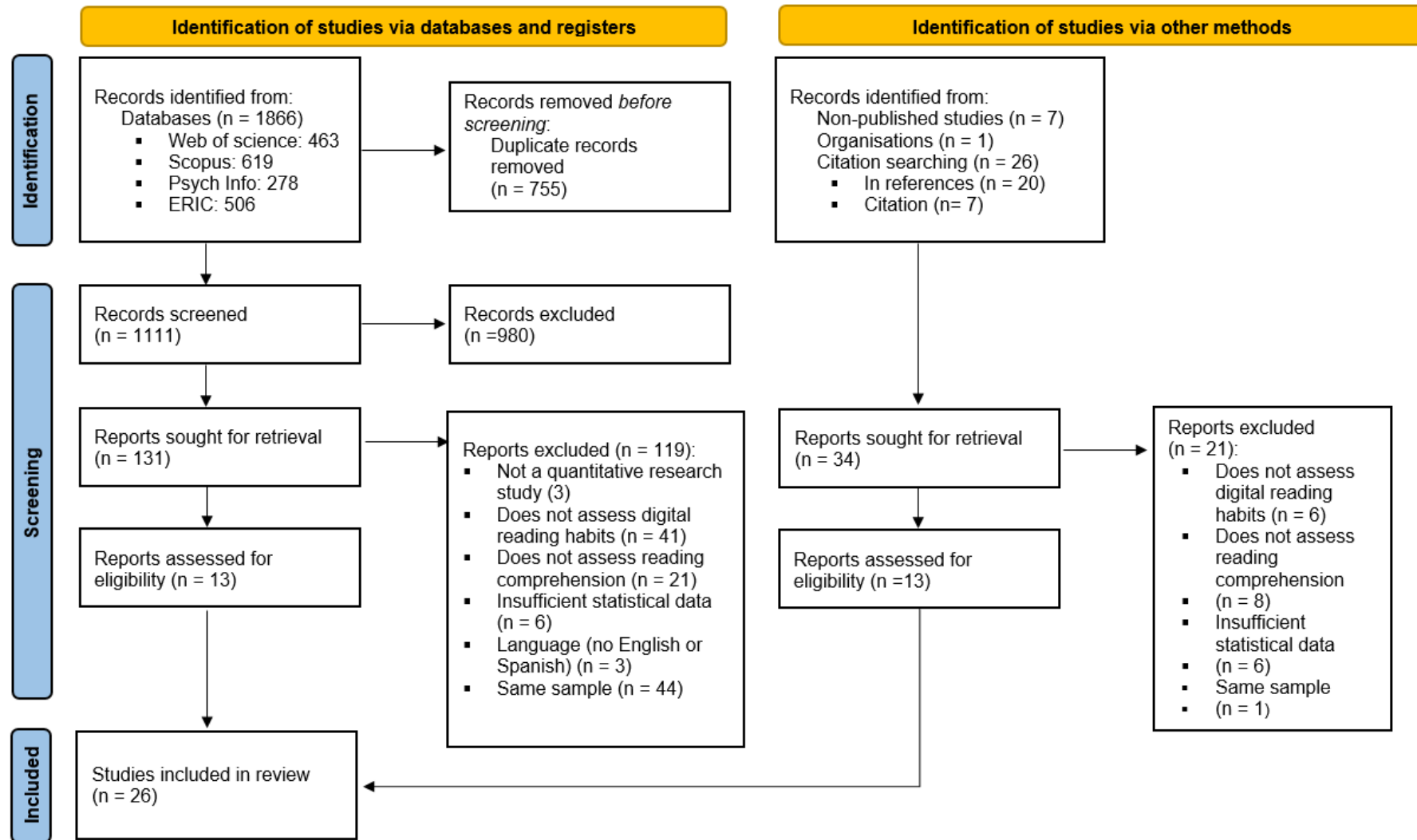
The search described above yielded 1,866 records. Following the criteria exposed, 26 studies were considered to perform the effect size calculus, including data from 469,564 individuals. A flow diagram based on PRISMA guidelines (Page et al., 2021) outlines this process in Figure 1.

Coding the studies

The first author and an independent research assistant coded all the reviewed articles selected randomly. Inter-rater reliability was adequate, showing a Cohen's kappa equal to .728 ($minimum_{Sampling}=.49$, $maximum_{Assessment RC}=.965$) for qualitative variables. Because the Kappa coefficient tends to be low, even though there may be

high agreement rates (Warrens, 2010), we also report the mean percentage of inter-rater agreement = 88.90% (minimum = 77.50%; maximum = 97.40%). For continuous variables, the intra-class correlation (95% CI) mean was .963 ($minimum_{Data\ collection\ year}=.89$, $maximum_{Mean\ age}=1$). Disagreements were discussed. For transparency and objectivity, a coding manual was developed and is available upon request. A descriptive overview of the studies included is given in the Results section and in the Appendix. Next, the variables coded across the studies are described.

PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources

**Figure 1.** PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources.

*Note: Selection for studies with the same sample was based on the criteria explained in the Effect size calculation section.

Pre-published version

Regarding substantive variables, we considered average age; percentage of females, and the corresponding educational stage for the participants at the moment the study was conducted. Educational stages were coded following the US system to provide a homogeneous categorization (1: Elementary school, from first to fifth grade; 2: Middle school, from sixth to eighth grades; 3: High school, from ninth to 12th grades; 4: Undergraduate students, higher education at university levels). To analyze methodological aspects, we included the following variables: total sample size and sampling procedure (0: non-probabilistic: convenient sampling vs. 1: probabilistic: random sample selection). Regarding the characteristics of the digital reading habits measurement, we coded the type of data collection instrument (in which 0: Self-report, refers to self-reported measurements such as filling out a questionnaire or survey and 1: Interview, refers to in-person interviews in which a third person gathered the information). Due to the variability among the reading habits questionnaires used in the primary studies, we coded the next variables regarding the questions and answers format (examples of these categories are provided in Table 1): temporal range (what period of time is being asked, 0: Non-specified, 1: Daily habits, 2: Weekly habits, 3: Mixed); answer format of the reading habits questionnaire (0: open vs. 1: closed); temporal classification of answers (0: Hours, 1: Frequency adverbs, 2: Times per day/week/month, 3: Mixed), number of answer-alternatives in the scale. Finally, to analyze the potential influence of the type of digital reading habits, we categorized the studies according to the items employed to assess habits. This moderator was labeled as “digital reading habits categories” and items were categorized as 0: Social-communicative reading (e.g., “Emails, online messages, social networks, blogs”; “Browsing the internet for fun, read or write emails to friends, write a blog or website, chatting online, participate in online forums, participate in social networks ”); 1: Linear-informative reading (e.g., “E-fiction, e-news, blogs and Bulletin Board Systems”; “e-books, online newspapers, websites on different topics, blog postings, and forums”); and 2: Mixed (when the proportion of items from categories 0 and 1 was equivalent or not clear, e.g., “Factual website or blog, online searching, text messages/e-mails, networking website, computer/console game, and Twitter”; “Electronic reading resources”). Addressing the differences across reading comprehension tests, we coded some variables regarding their characteristics: the media in which participants took the test (0: Print, 1: Digital-based); the type of assessment (0: Multiple choice questions, 1:

Open-ended questions, 2: CLOZE type, 3: Mixed); type of reading comprehension dimension assessed (0: Inferential, 1: Literal, 2: Mixed); genre of the text or texts used (0: Narrative, 1: Expository, 2: Mixed). Finally, regarding the reliability of the instruments employed, we considered the reliability coefficients of the reading comprehension test and the reading habits questionnaire. Lastly, we coded two extrinsic variables: the data collection year (when the study was conducted) and the publication status (whether the report was published or not).

Effect size calculation

The effect size used to measure the relationship between leisure digital reading habits and reading comprehension scores was the Pearson correlation coefficient (r) performing multilevel random-effects analysis. If the coefficients were calculated for each item of the reading habit questionnaire separately, we performed an average score by merging this information into one variable (*digital reading habits*).

If the study did not include these correlational analyses, we used two strategies to perform the calculation. First, we tried to use the descriptive data that allowed us to calculate the effect size: mean, standard deviations and sample size for both variables (e.g., the score in reading comprehension according to each reading frequency option). Second, when none of these data were reported, we looked for multiple regression analysis. In this specific case, we could impute the effect size using the following formula:

$$r = \beta + .05\lambda$$

This formula can be applied only when beta is standardized and when the value is inside a range of - 0.5 to + 0.5. The λ variable is an indicator that equals 1 when β is non negative and 0 when β is negative. This procedure was developed by Peterson and Brown (2005) who proposed and evidenced an appropriate approach for imputing r -based meta-analytic effect sizes in behavioral studies using knowledge of corresponding beta coefficients. Ultimately, if any of these strategies could not be carried out, or necessary information was missing to perform other imputation strategies, we wrote to the authors asking for data.

To maintain the effect size independence across different studies, we followed specific selection criteria to identify the reports that best fitted our objectives and, additionally, covered a wider population range. Some pre-selected studies worked with

the same database. Mostly, this was the case of PISA and PIRLS databases (Program for International Student Assessment, n.d., Progress in International Reading Literacy Study, n.d.). Therefore, to carry out the selection, the following criteria were applied by relevance order (a comparison table is available at the Appendix, Table A):

- 1) Digital reading habits. We prioritized those studies that included more items referred to digital reading habits. This criterion was applied mostly in the selection of PISA studies, as different authors included different items from the same sample.
- 2) Sample size. We prioritized those studies that included a larger sample (usually from different countries). In the case of PISA studies, we discarded the studies that used information from a limited number of countries and were already addressed in other broader studies. For example, a study using information from 10 countries would be more relevant than one using 4.
- 3) Sample independence. We prioritized those studies that worked with more independent samples (e.g., data from different countries or age-groups).

Regarding longitudinal studies, which included data from the same participants across several years, we decided to take the cross-sectional information from the first year in which participants were assessed. In this way, we guaranteed no dragging effect.

Statistical analyses

Pearson correlation coefficients were transformed into the Fisher's Z metric to normalize distributions and stabilize variances. To facilitate the interpretation of the results, they were transformed back into correlation coefficients. A three-level random-effect meta-analysis was applied to address the fact that some effect sizes were nested within studies. This model suits our data best because several of the primary studies analyzed contained information from independent sub-samples (Cheung, 2019). Instead of aggregating the information for each sub-sample, with this multilevel analysis we were able to work with individual effect sizes while controlling dependence at the study-level, increasing statistical power as well (Harrer et al., 2021). In this model, the first level represents a within-effect size model, the second level shows variation between the effect sizes within the same study, the third one represents variation across studies (Van den Noortgate et al., 2013). Despite the fact that the participant level only contributed to one effect size (the sampling errors are conditionally independent), the

non-independence could be introduced to the nested structure of the effect size. With the three-level random-effect analysis we could analyze the between-study heterogeneity that would be undetected in aggregated data. We opted for this approach, instead of nesting studies from the same research teams, because variation of sample characteristics and instruments used tended to be more diverse across studies than across teams. Thus, we calculated cluster-robust standard errors, statistical tests, and confidence intervals for our estimates by using the “CR2” method (Tipton & Pustejovsky, 2015). The method used to estimate the model parameters was restricted maximum-likelihood (REML). The inverse variance method was used to weigh each effect size. An average effect size and a 95% CI was calculated. Furthermore, outliers’ cases were assessed with standardized deleted residuals larger than 1.96 (Viechtbauer, 2021; Viechtbauer & Cheung, 2010).

Cochran's Q statistic was used to assess the presence of heterogeneity. The multilevel I^2 statistic was also considered to estimate the proportion of the variance in observed outcomes that reflects variation in true effect sizes rather than sampling error, with $I^2_{\text{Level 2}}$ and $I^2_{\text{Level 3}}$ representing within- and between-study heterogeneity, respectively. Moreover, a 95% prediction interval around the main effect size was calculated to know how the true effect varied across populations (Borenstein, 2019; Cheung, 2014).

To explore heterogeneity, the analyses of moderator variables were performed by using multilevel meta-regression models with robust variance estimation method for continuous and categorical moderators.

Sensitivity analyses were carried out to assess the robustness of the results. First, the average effect size was calculated for the whole sample of studies and for the sample without the outlier (Gumus & Atalmis, 2011). Second, we ran the same meta-regression analysis for the moderators including the outlier. Third, due to the big variability of sample sizes, the average effect size was compared between those studies which presented extra-large sample sizes and the rest of the studies. Publication bias was assessed using Egger's linear regression (Card, 2012), and applying multilevel meta-regression models to evaluate whether publication status moderated the effect size.

The statistical analyses were performed using R 4.2.1 version with Metafor (Viechtbauer, 2010) and clubSandwich (Pustejovsky, 2020) packages.

Results

Descriptive characteristics of the studies

Twenty-six studies fulfilled the selection criteria. Because six of these studies contained independent subsamples, we could analyze a total of 40 independent effect sizes with 469,564 participants. The main characteristics of the studies are described in Table 2.

In regard to sociodemographic variables, the average age was 15.9 years (range: 9.9-72.5) and the mean percentage of female participants was 54.84% (range: 41.5-92.0). The majority of participants were middle school (35.89%) and high school (35.89%) students. Regarding the methodological variables, 66.67% of effect sizes were from studies that recruited the sample through a non-probability sampling method. The most common data extraction method for measuring digital reading habits was the questionnaire (97.37%) with restricted questions (97.7%) (as only one study did not use questionnaires and only one measured the reading habits frequency with open answer format, these variables were not included in the moderator analysis). On average, reliability of the reading comprehension tests used (Cronbach $\alpha = .77$) was higher than that of the digital reading habit questionnaires ($\alpha = .59$). Finally, in the case of extrinsic variables, 56.41% of effect sizes were from published research.

The mean effect size

To start, an exploratory analysis was conducted in order to assess influential cases and normality assumption before analyzing the average effect size. A particular effect size was identified as an influential outlier (Gumus & Atalmis, 2011; standardized deleted residuals larger than 3.42), and accordingly it was not used to estimate the average correlation. Therefore, the analyses presented include 39 effect sizes.

The overall association between leisure digital reading habits and reading comprehension was significant and positive, although the effect was rather small ($r = .055$; 95% CI: .003, .107; $p = .03$; $k = 39$, $n = 25$). Therefore, on average, people that engage in more frequent leisure digital reading activities tend to score higher in reading comprehension tasks.

In reference to the variability of the effect sizes, the heterogeneity Q statistic reached statistical significance, $Q(38) = 2591.93$, $p < .001$. The I^2 statistic was 98.9, that is, 98.9 of the variance in the observed effects reflected variance in true effects rather than sampling error. The majority of true variance comes from between-study variance (68.5%), versus within-study variance (30.5%). Furthermore, the prediction interval was -.18 to .29. This range is usually wider than the confidence interval, particularly when there is great heterogeneity in the data (Al Amer & Lin, 2021), as is the case in our meta-analysis. We would expect that the true effect size of all comparable populations would fall in this range in 95% of all populations.

Analyzing moderator variables

The heterogeneity found led to an analysis of the influence of potential moderator variables. Table 3 and 4 show the results of applying multilevel meta-regression analyses with robust variance estimation to assess the influence of moderator variables on the effect sizes for qualitative and quantitative moderators, respectively.

With regard to the sociodemographic variables, educational stage was a significant moderator ($p = .047$; $R^2 = .57$). We performed pairwise comparisons across the levels and two comparisons were significant: the difference between middle and high school students ($p = .021$) and between primary and university students ($p = .030$). The results showed that the effect size for high school students was positive and higher ($r = .085$) than the effect size obtained in middle school students ($r = -.025$). In addition, the mean age of the participants exhibited a positive relationship with the effect sizes, although this difference was not statistically significant ($p = .175$). As most of the samples included student populations (97.40%), we decided to narrow down the selection. In this regard, we carried out a second analysis in which we included studies whose participants were students, excluding only 2 reports (Champley et al., 2008; Smith & Smith; 2010). The regression analysis turned out to be statistically significant only when including primary and secondary school students ($p = .03$; $R^2 = .38$) and revealed higher effect sizes as the mean age of the students increased ($\beta = .012$). Finally, the percentage of female participants did not reach a statistically significant relationship with the effect sizes ($p = .181$; $R^2 = .00$).

About the methodological variables, the multilevel analysis did not identify any variable as a significant predictor. Contrary to our expectations, no differences were

observed regarding the type of digital reading habit and reading comprehension ($p = .18$): social-communicative habits ($r = .03$), linear-informative habits ($r = .12$).

Regarding the extrinsic variables, differences have not been found in the magnitude of the effect sizes in regard to published status ($p = .53$; $R^2 = .00$). This was also confirmed by the publication bias analysis (see the sensitivity section). Lastly, contrary to our expectations, the data collection year was not significant as a predictor ($p = .121$).

Sensitivity and bias analysis

For the sensitivity analysis, we calculated the average effect size for the studies including the outlier (see Table 5). When comparing the calculations without the outlier, we observe that the average value is slightly higher and also significant. The main difference is found in between-cluster variance, being almost half when removing the outlier. In this case, σ^2 is equivalent to the between-study heterogeneity variance (τ^2) in a conventional meta-analysis (since clusters represent studies in our model). Second, we performed multilevel meta-regression analysis for moderator variables including the outlier. The significance of the moderators remained the same except for the data collection year which turned out significant ($b_{with\ outlier} = -.014$; 95% CI : $-.27, -.001$; $p = .043$).

In order to assess whether publication bias might be a threat against our meta-analytic results, two techniques were applied. The Egger test of the intercept did not reach a statistically significant result ($b = -.039$; 95% CI : $-1.39, .61$; $p = .427$). In addition, multilevel meta-regression analyses with robust variance estimation revealed that the mean effect sizes from published versus unpublished studies were not statistically different ($p = .525$). Thus, these results suggested that there was no publication bias.

Lastly, five studies were categorized as having extra-large samples ($N > 10.000$): Kong et al., (2022), Park & Weng (2020), Petko et al., (2017), Smith & Smith (2010), and Wu & Peng (2017). This criterion was followed based on a recent meta-analysis (Lam et al., 2021). We did not find statistically significant differences between studies with extra-large samples and the rest of studies ($p = .153$). This means that the relationship between digital reading habits and reading comprehension from studies with small and large sample sizes are not statistically different.

Discussion

This meta-analysis sought to shed light on the inconclusive relationship between leisure digital reading habits and reading comprehension and to identify potential moderator variables for such relationship. By conducting multilevel analysis, this work provides robust information controlling for intra-studies dependencies. Still, we found great heterogeneity among studies, calling for a cautious interpretation of the data.

Regarding our main objective, the results evidenced a positive and significant association between digital reading habits and reading comprehension. The effect size is rather small, especially when compared with previous meta-analytical studies that have evidenced medium to strong positive associations between print-reading habits and comprehension (Mol & Bus, 2011). Regarding moderator analysis, students' academic stage played a significant role on the association between reading comprehension and digital reading habits: the relation is negative in younger populations (i.e., primary and middle school students), and becomes positive in older populations (i.e., secondary school and undergraduate students). All in all, these results allow us to untangle the relationship between reading habits and reading comprehension when it comes to the digital context, revealing that leisure digital reading habits do not pay off as much as traditional print reading.

Why is the relationship between reading habits and reading comprehension so different when it comes to different media? At this point, we can only speculate about those divergent patterns. Our results do not support the predictions from the displacement hypothesis, as the current change from print to digital reading habits decreased dramatically the association between habits and reading comprehension. Accordingly, we cannot just assume that the societal move towards digital reading habits mobilizes and spurs the same comprehension resources as previously reported for print reading habits. An alternative explanation could rely on the *shallowing hypothesis* (Annisette & Lafreniere, 2017; Carr, 2010), which emphasizes that the superficial way in which people tend to interact with new forms of digital reading may not support readers' engagement or further development of higher-order cognitive skills. The frequent characteristics of digital reading texts, such as short length and fast-paced stimuli, emphasize shallow processing (Baron, 2021; Ra et al., 2018). Additionally, the lower linguistic quality of digital texts may prevent students from being exposed to

academic models typically found in traditional printed texts (Snow, 2010). Future research should continue exploring those issues.

Digital reading habits and the development of reading comprehension

Our analysis contributes to the ongoing debate on the influence of digital technologies on the development of reading comprehension (Wolf, 2018). Previous longitudinal research regarding the relation between reading habits and comprehension has already highlighted such relation varies across the life span (Locher & Pfof, 2020; Torppa et al., 2019; van Bergen et al., 2020). Nevertheless, there is still no clear developmental pattern for such relation. Our exploratory analysis indicated a significant negative relationship between leisure digital reading habits and reading comprehension at earlier educational stages such as primary and middle school, compared to high school, in which the relationship turned out to be positive, as was the case with undergraduate students. This pattern concurs with recent evidence showing a stronger negative association between the frequency of use of digital devices in the classroom and reading comprehension in US representative samples of fourth grade, as compared to eighth grade (Salmerón et al., 2023). Nevertheless, due to the correlational nature of the original studies analyzed, we can only speculate about the underlying reasons for those differences. Such results could be interpreted in light of potential early sensitive developmental periods. Great exposition to digital reading activities, which frequently contain low linguistic quality texts, may detract early readers from building a strong reading foundational base. For example, early readers engaging in frequent digital reading may learn less academic vocabulary and syntax (Snow, 2010), or develop to a lesser extent the ability to keep focused on a task (Koolstra et al., 1997), in a period when they are critical to shift from learning to read to reading to learn. Future research should test those and other potential causal mechanisms for the developmental variations in the association between digital reading habits and reading comprehension.

Beyond the developmental aspects, we expected that the type of digital reading habits would contribute differently to reading comprehension (Naumann & Sälzer, 2017). Our moderator analysis regarding this factor did not reach statistical significance, a pattern that does not align with previous research showing larger effect sizes for linear-communicative digital reading habits than for social-communicative ones (Duncan et al., 2016; McGeown et al., 2016; Pfof et al., 2013). Future research may

clarify the nature of those effects by analyzing the combined relation between types of digital reading habits and print reading ones, in light of the *displacement hypothesis*. Combining digital linear-informative reading habits with print reading habits may be particularly beneficial. In the same line, research is needed to identify how different types of digital reading habits influence comprehension across the life span. This type of research would disentangle the cause of the negative associations between digital reading habits and reading comprehension observed for younger students, which may potentially be restricted to social-communicative habits.

Limitations and future research

There are some limitations in this meta-analysis that should be discussed. Three relevant limitations are related to the existing datasets. First, we found several studies that belonged to the same research teams, which may add unexpected bias to the results. As explained in the method section, we tried to control as much as possible the dependency effects between each effect size by using a multilevel approach that controlled for the level of student, sub-sample, and study. As this research field continues to expand, future analyses would be able to include primary studies from a larger range of teams. Second, most of the primary studies used a cross-sectional design, so we relied on correlational coefficients that cannot inform us about causal effects. A larger number of longitudinal studies is needed in order to provide accurate estimates about causal relations between habits and comprehension. Third, most of the primary studies measured reading habits via self-reports with closed-answer options; thus the information could be biased and tied to human-measurement errors.

Moreover, we only included leisure digital reading habits. Therefore, research focusing on other reading purposes, like academic reading, or exploring digital and print reading habits in combination, would be relevant for this matter. In addition, the scope of the search was narrowed down to assessments in native language and did not include studies that were focused on readers with learning difficulties or physical impairments. From the lenses of the Matthew effects on reading, one could expect that the relationship between digital reading habits and reading comprehension may be smaller or even negative in students with learning difficulties. Future research could clarify this issue by carrying studies focusing on this type of readers. Referring to our dependent variable, the studies included in this meta-analysis focused on traditional text reading

comprehension (i.e., linear texts). Future research should explore the associations between digital reading habits and digital reading comprehension (i.e., hypertext, Internet documents). Finally, from a methodological perspective, we highlight that the nature of our data is aggregated, meaning that we have worked with means which do not reflect individual differences inside the studies. In addition, pseudo R^2 coefficient used in multilevel meta-regression analysis cannot be considered fully stable when the number of introduced studies is below 40 (Van den Noortgate et al., 2013). For this reason, we recommend taking the proportion of explained variance with caution.

As for the difficulties encountered, we highlight the heterogeneity of the scales and instruments used across studies, both regarding digital reading habits and reading comprehension measurements. Creating a standardized scale to assess digital reading habits may help to set a common ground for new studies. However, given the evolving nature of the tasks offered by new digital tools and services, such scale may be difficult to create. That is why, when proposing future investigations, research in this field could benefit from using methods such as questionnaires to address common reading habits and, additionally, exploring qualitative dimensions such as the length of reading, motivation, and attitudes towards reading (Kaye et al., 2020). Moreover, triangulation with other data collection techniques should include more objective measures (such as phone use time-trackers or observational techniques), as a recent systematic review has evidenced that the correspondence between self-reported and logged social media use is far from perfect (Parry et al., 2021).

Conclusion

While the use of digital devices in education is an undeniable reality nowadays, our results emphasize the need to increase the awareness of how students interact with them and how this relates to important educational factors such as text comprehension. Based on our results, we cannot just assume that leisure reading times would be beneficial for developing readers, regardless of the medium used. The notable difference between the effect sizes of previous meta-analysis regarding print reading habits and the present work leads us to search for further explanations. Based on our moderator analysis, we cautiously recommend educators and practitioners to pay more attention to how different digital reading habits can play a role in the reading comprehension of younger readers. Because the heterogeneity found across effect sizes

is rather big, we consider this meta-analysis only as a first step into exploring the relationship between reading comprehension and digital reading habits.

Acknowledgements

Special thanks to Helge Strømsø for his insight in an earlier version of the manuscript, to Anastasia Karagiorgi and Mario Romero for their collaboration in the analyses and to the authors that took the time to answer our request and provided the information needed to conduct the analysis.

References

References included in the systematic review and meta-analysis are marked with an asterisk.

- Al Amer, F. M., & Lin, L. (2021). Empirical assessment of prediction intervals in Cochrane meta-analyses. *European Journal of Clinical Investigation*, 51(7), e13524. <https://doi.org/10.1111/eci.13524>
- *Altamura, L. (2019). Digital reading habits and reading comprehension [Unpublished Bachelor Dissertation, University of Valencia]. <https://go.uv.es/D4FyR5I>
- *Amiama-Espallat, C., & Mayor-Ruiz, C. (2017). Digital reading and reading competence—the influence in the Z generation from the Dominican Republic. *Comunicar. Media Education Research Journal*, 52, 105-114. <https://doi.org/10.3916/C52-2017-10>
- 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>
- Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology*, 72(2), 261-278. <https://doi.org/10.1348/000709902158883>
- Bann, C. M., Bayen, U. J., McCormack, L. A., & Uhrig, J. D. (2006). Effects of reading habits, reading comprehension, and memory beliefs on older adults' knowledge about Medicare. *Journal of Applied Gerontology*, 25(1), 49-64. <https://doi.org/10.1177/0733464805283627>
- Baron, N. S. (2021). Know what? How digital technologies undermine learning and remembering. *Journal of Pragmatics*, 175, 27-37. <https://doi.org/10.1016/j.pragma.2021.01.011>

- Baron, N. S., Calixte, R. M., & Havewala, M. (2017). The persistence of print among university students: An exploratory study. *Telematics and Informatics*, 34(5), 590-604. <https://doi.org/10.1016/j.tele.2016.11.008>
- Borenstein, M. (2019). *Common Mistakes in Meta-Analysis and How to Avoid Them*. Biostat, Inc.
- Botello-Peñaloza, H.A. & López-Alba A. (2014). La influencia de las TIC en el desempeño académico: Evidencia de la prueba PIRLS en Colombia 2011 [ICT influence in academic performance: 2011 PIRLS evidence from Colombia]. *Revista Academia y Virtualidad*, 7(2), 15-26. <https://doi.org/10.18359/ravi.315>
- Bråten, I., & Braasch, J. L. G. (2018). The role of conflict in multiple source use. In J.L.G. Braasch., I. Bråten., & M.T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 184-201). (2018). Routledge. <https://doi.org/10.4324/9781315627496>
- Caccia, M., Giorgetti, M., Toraldo, A., Molteni, M., Sarti, D., Vernice, M., & Lorusso, M. L. (2019). ORCA. IT: A new web-based tool for assessing online reading, search and comprehension abilities in students reveals effects of gender, school type and reading ability. *Frontiers in Psychology*, 10, 2433. <https://doi.org/10.3389/fpsyg.2019.02433>
- Card, N. A. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.
- Carr, N. (2010). *The shallows: How the internet is changing the way we think, read and remember*. Atlantic Books Ltd.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- *Champley, J., Scherz, J.W., Apel, K. and Burda, A.N. (2008). A preliminary analysis of reading materials and strategies used by older adults. *Communication Disorders Quarterly*, 29(3), 131-140. <http://dx.doi.org/10.1177/1525740108315365>
- Chen, C. M., & Chen, F. Y. (2014). Enhancing digital reading performance with a collaborative reading annotation system. *Computers & Education*, 77, 67-81. <https://doi.org/10.1016/j.compedu.2014.04.010>

- *Chen, S. Y., & Fang, S. P. (2013). Developing a Chinese version of an author recognition test for college students in Taiwan. *Journal of Research in Reading*, 38(4), 344-360. <https://doi.org/10.1111/1467-9817.12018>
- Chen, J.L. & Su, Y.F. (2010). *The effects of different reading strategy instructions on low reading ability college students' reading strategy performance: Evidences from think aloud data* [Paper presentation]. 20th Annual meeting of the Society for Text & Discourse, Chicago, IL.
- Cheung, M. W. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, 19(2), 211-229. <https://doi.org/10.1037/a0032968>
- Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review*, 29(4), 387-396. <https://doi.org/10.1007/s11065-019-09415-6>
- Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, 42(2), 288-325. <https://doi.org/10.1111/1467-9817.12269>
- Clinton-Lisell, V. (2021). Stop multitasking and just read: Meta-analyses of multitasking's effects on reading performance and reading time. *Journal of Research in Reading*, 44(4), 787-816. <https://doi.org/10.1111/1467-9817.12372>
- Clinton- Lisell, V., Seipel, B., Gilpin, S., & Litzinger, C. (2023). Interactive features of E-texts' effects on learning: A systematic review and meta-analysis. *Interactive Learning Environments*, 31(6), 3728-3743. <https://doi.org/10.1080/10494820.2021.1943453>
- Coiro, J. (2021). Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Reading Research Quarterly*, 56(1), 9– 31. <https://doi.org/10.1002/rrq.302>
- Cuetos, E., Rodríguez, B., Ruano, E. & Arribas, D. (2014). PROLEC-R. *Batería de evaluación de los procesos lectores, revisada (5a ed.)* [Battery for the reading process assessment, reviewed (5th ed.)]. TEA Ediciones.
- Cuetos, F., Arribas, D. & Ramos, J.L. (2016). PROLEC-SE-R. *Batería de evaluación de los procesos lectores en secundaria y bachillerato, revisada* [Battery for the reading process assessment in middle and high school, reviewed]. TEA Ediciones.

- Delgado, P., & Salmerón, L. (2021). The inattentive on-screen reading: Reading medium affects attention and reading comprehension under time pressure. *Learning and Instruction, 71*, 101396. <https://doi.org/10.1016/j.learninstruc.2020.101396>
- 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>
- Duchêne, M.C.A. (2010). *Protocole Emilie – Evaluation de la compréhension de textes chez les collégiens [Emilie protocol – Text comprehension assessment for secondary school students]*. Ortho Edition.
- *Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology, 107*(2), 209-238. <https://doi.org/10.1111/bjop.12134>
- Dunn, L.M., & Markwardt, F.C. (1970). *Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Services.
- Forzani, E. & Leu, D.J. (2012). New literacies for new learners: The need for digital technologies in primary classrooms. *The Educational Forum, 76*(4), 421-424. <https://doi.org/10.1080/00131725.2012.708623>
- Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A Comparison of children's reading on paper versus screen: a meta-analysis. *Review of Educational Research, 91*(4), 483–517. <https://doi.org/10.3102/0034654321998074>
- García, V. (2019). Construcción de mapas conceptuales navegables y comprensión lectora: análisis de procesos y diseño de instrucción [Navigable concept maps and Reading comprehension: process analysis and instruction design] [Doctoral dissertation, University of Valencia]. <https://roderic.uv.es/handle/10550/70352>
- *Gumus, S., & Atalmis, E. H. (2011). Exploring the relationship between purpose of computer usage and reading skills of Turkish students: evidence from PISA 2006. *Turkish Online Journal of Educational Technology-TOJET, 10*(3), 129-140. <https://bit.ly/2TTE7Z8>

- Harrer, M., Cuijpers, P., Furukawa, T.A., & Ebert, D.D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide*. Boca Raton, FL and London: Chapman & Hall/CRC Press. ISBN 978-0-367-61007-4.
- *Hsu, C. T., Clariana, R., Schloss, B., & Li, P. (2019). Neurocognitive signatures of naturalistic reading of scientific texts: a fixation-related fMRI study. *Scientific Reports*, 9(1), 1-16. <https://doi.org/10.1038/s41598-019-47176-7>
- Islam, M. I., Biswas, R. K., & Khanam, R. (2020). Effect of internet use and electronic game-play on academic performance of Australian children. *Scientific Reports*, 10(1), 21727. <https://doi.org/10.1038/s41598-020-78916-9>
- Jackson, L. A., von Eye, A., Fitzgerald, H. E., Witt, E. A., & Zhao, Y. (2011). Internet use, videogame playing and cell phone use as predictors of children's body mass index (BMI), body weight, academic performance, and social and overall self-esteem. *Computers in Human Behavior*, 27(1), 599–604. <https://doi.org/10.1016/j.chb.2010.10.019>
- *Johnson, G. M. (2013). Traditional literacy skills and internet use among 8-to 12-year-old children. *Reading Psychology*, 34(5), 486-506. <https://doi.org/10.1080/02702711.2012.658144>
- *Kannianen, L., Kiili, C., Tolvanen, A., Utriainen, J., Aro, M., Leu, D.J., & Leppänen, P.H.T. (2022). Online research and comprehension performance profiles among sixth-grade students, including those with reading difficulties and/or attention and executive function difficulties. *Reading Research Quarterly*, 57(4), 1213–1235. <https://doi.org/10.1002/rrq.463>
- Kaye, L., Orben, A., Ellis, D. A., Hunter, S. C., & Houghton, S. (2020). The conceptual and methodological mayhem of “screen-time”. *International Journal of Environmental Research and Public Health*, 17(10), 3661. <https://doi.org/10.3390/ijerph17103661>
- 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>
- *Kong, Y., Seo, Y. S., & Zhai, L. (2022). ICT and digital reading achievement: a cross-national comparison using PISA 2018 Data. *International Journal of Educational Research*, 111, 101912. <https://doi.org/10.1016/j.ijer.2021.101912>

- Koolstra, C.M., van der Voort, T.H.A. & van der Kamp, L.J.T. (1997), Television's Impact on Children's Reading Comprehension and Decoding Skills: A 3-Year Panel Study. *Reading Research Quarterly*, 32(2), 128-152.
<https://doi.org/10.1598/RRQ.32.2.1>
- Lam, K. K. L., & Zhou, M. (2022). Grit and academic achievement: A comparative cross-cultural meta-analysis. *Journal of Educational Psychology*, 114(3), 597–621. <http://dx.doi.org/10.1037/edu0000699>
- Lerikkanen, M.K., Eklund, K., Löytynoja, H., Aro, M., Poikkeus, A.M. (2018). *YKÄ—Luku- ja kirjoitustaidon arviointimenetelmä yläkouluun [YKÄ—Reading test for lower secondary school]*. Jyväskylä, Finland: Niilo Mäki Instituutti
- Leslie, L., & Caldwell, J. S. (2001). *Qualitative Reading Inventory-3* (3rd ed.). New York: Harper Collins.
- Leu, D.J., Gregory McVerry, J., Ian O'Byrne, W., Kiili, C., Zawilinski, L., Everett-Cacopardo, H., Kennedy, C. and Forzani, E. (2011). The new literacies of online reading comprehension: Expanding the literacy and learning curriculum. *Journal of Adolescent & Adult literacy*, 55(1), 5-14. <https://doi.org/10.1598/JAAL.55.1.1>
- Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly*, 50(1), 37-59.
<https://doi.org/10.1002/rrq.85>
- Liu, Y., & Gu, X. (2020). Media multitasking, attention, and comprehension: A deep investigation into fragmented reading. *Educational Technology Research and Development*, 68(1), 67-87. <https://doi.org/10.1007/s11423-019-09667-2>
- Liu, D., Kirschner, P. A., & Karpinski, A. C. (2017). A meta-analysis of the relationship of academic performance and Social Network Site use among adolescents and young adults. *Computers in Human Behavior*, 77, 148-157.
<https://doi.org/10.1016/j.chb.2017.08.039>
- Llorens, A. C., Gil, L., Vidal-Abarca, E., Martínez, T., Mañá, A., & Gilabert, R. (2011). Prueba de competencia lectora para educación secundaria (CompLEC) [Reading literacy text for secondary education]. *Psicothema*, 23, 808-817.
<http://www.redalyc.org/articulo.oa?id=72722232043>

- Locher, F., & Pfof, M. (2020). The relation between time spent reading and reading comprehension throughout the life course. *Journal of Research in Reading*, 43(1), 57-77. <https://doi.org/10.1111/1467-9817.12289>
- Macmillan Test Unit. (2000). *Group Reading Test II*, 6 – 14. Windsor: NFER-Nelson.
- *Macedo-Rouet, M., Salmerón, L., Ros, C., Pérez, A., Stadtler, M., & Rouet, J. F. (2020). 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), 101-138. <https://doi.org/10.1080/02103702.2019.1690849>
- Mak, M. T. F., Wang, M., & Chu, K. W. S. (2019). Effects of a gamified learning platform on elementary school students' flow experiences in leisure reading. *Proceedings of the Association for Information Science and Technology*, 56(1), 454-458. <https://doi.org/10.1002/pra2.75>
- Mangen, A. (2008). Hypertext fiction reading: haptics and immersion. *Journal of Research in Reading*, 31(4), 404-419. <https://doi.org/10.1111/j.1467-9817.2008.00380.x>
- 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>.
- Martínez, T., Vidal-Abarca, E., Sellés, P., & Gilabert, R. (2008). Evaluación de las estrategias y procesos de comprensión: el test de procesos de comprensión. [Assessment of strategies and reading comprehension processes: the reading comprehension processes test]. *Infancia y Aprendizaje*, 31(3), 319-332. <https://doi.org/10.1174/021037008785702956>
- *McGeown, S. P., Osborne, C., Warhurst, A., Norgate, R., & Duncan, L. G. (2016). Understanding children's reading activities: Reading motivation, skill and child characteristics as predictors. *Journal of Research in Reading*, 39(1), 109-125. <https://doi.org/10.1111/1467-9817.12060>
- Merga, M. K., & Roni, S. M. (2017). The influence of access to eReaders, computers and mobile phones on children's book reading frequency. *Computers & Education*, 109, 187–196. <https://doi.org/10.1016/j.compedu.2017.02.016>

- *Mežek, Š. (2013). Multilingual reading proficiency in an emerging parallel-language environment. *Journal of English for Academic Purposes*, 12(3), 166-179.
<https://doi.org/10.1016/J.JEAP.2013.02.001>
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296. <https://psycnet.apa.org/doi/10.1037/a0021890>
- National Assessment Program Literacy and Numeracy (NAPLAN). (2013). *NAPLAN Achievement in Reading, Persuasive Writing, Language Conventions and Numeracy: National Report for 2013*. ACARA, Sydney.
- National Center for Educational Statistics (NCES). (n.d.). *What is NAAL?*
<http://nces.ed.gov/naal/index.asp>
- Naumann, J., & Sälzer, C. (2017). Digital reading proficiency in German 15-year olds: Evidence from PISA 2012. *Zeitschrift für Erziehungswissenschaft*, 20(4), 585-603. <https://doi.org/10.1007/s11618-017-0758-y>
- Neuman, S. B. (1988). The displacement effect: Assessing the relation between television viewing and reading performance. *Reading Research Quarterly*, 23(4), 414-440. <https://doi.org/10.2307/747641>
- Nikkelen, S. W. C., Valkenburg, P. M., Huizinga, M., & Bushman, B. J. (2014). Media use and ADHD-related behaviors in children and adolescents: A meta-analysis. *Developmental Psychology*, 50(9), 2228–2241. <https://doi.org/10.1037/a0037318>
- *Norris, T. (2010). *Adolescent academic achievement, bullying behavior, and the frequency of Internet use* [Doctoral dissertation, Kent State University].
ProQuest Dissertations and Theses Global.
<https://www.proquest.com/docview/193947450>
- Organization for Economic Cooperation and Development (OECD). (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*.
<https://doi.org/10.1787/9789264026407-en>.
- Organization for Economic Cooperation and Development (OECD). (2010). *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. <https://doi.org/10.1787/9789264062658-en>.
- Organization for Economic Cooperation and Development (OECD). (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science,*

Problem Solving and Financial Literacy.

<http://dx.doi.org/10.1787/9789264190511-en>

Organization for Economic Cooperation and Development (OECD). (2017). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving.*

<https://doi.org/10.1787/9789264281820-en>.

Organization for Economic Cooperation and Development (OECD). (2019). *PISA 2018 Assessment and Analytical Framework Mathematics, Reading, Science, Problem Solving and Financial Literacy.* <https://doi.org/10.1787/b25efab8-en>

Organization for Economic Cooperation and Development (OECD). (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do, PISA.* <https://doi.org/10.1787/5f07c754-en>.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.)*, 372, 71. <https://doi.org/10.1136/bmj.n71>

*Park, S., & Weng, W. (2020). The relationship between ICT-related factors and student academic achievement and the moderating effect of country economic index across 39 countries. *Educational Technology & Society*, 23(3), 1-15. <https://bit.ly/3rVyTc2>

Parry, D. A., Davidson, B. I., Sewall, C. J., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 5, 1535-1547. <https://doi.org/10.1038/s41562-021-01117-5>

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, 11(4), 357-383. <https://doi.org/10.1080/10888430701530730>

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, 90(1), 175. <https://doi.apa.org/doi/10.1037/0021-9010.90.1.175>

*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(8), 1070–1091. <https://doi.org/10.1177/0735633116649373>
- Pfost, M., Dörfler, T., & Artelt, C. (2012). Reading competence development of poor readers in a German elementary school sample. An empirical examination of the Matthew effect model. *Journal of Research in Reading*, 35(4), 411–426. <https://doi.org/10.1111/j.1467-9817.2010.01478.x>.
- *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>
- 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(2), 203–244. <https://doi.org/10.3102/0034654313509492>
- Program for International Student Assessment (PISA). (n.d.). *PISA Database*. <https://www.oecd.org/pisa/data/>
- Progress in International Reading Literacy Study (PIRLS). (n.d.). *Data Files*. <https://nces.ed.gov/surveys/pirls/datafiles.asp>
- Pustejovsky, J. E. (2020). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections (0.4.2) [R package]. <https://github.com/jepusto/clubSandwich>
- Ra, C. K., Cho, J., Stone, M. D., De La Cerda, J., Goldenson, N. I., Moroney, E., Tung, I., Lee, S. S., & Leventhal, A. M. (2018). Association of digital media use with subsequent symptoms of attention-deficit/hyperactivity disorder among adolescents. *JAMA*, 320(3), 255–263. <https://doi.org/10.1001/jama.2018.8931>
- Salmerón, L., Altamura, L., Blanco-Gandia, M. C., Montagud, S., & Vargas, C. (2022). Validation of a reading comprehension efficiency test for Spanish university students. *Psicologica* 43(2), e14809. <https://doi.org/10.20350/digitalCSIC/14809>
- *Salmerón, L., García, A., & Vidal-Abarca, E. (2011). [Unpublished raw data]. ERI Lectura, University of Valencia.
- *Salmerón, L., García, A., & Vidal-Abarca, E. (2018). The development of adolescents' comprehension-based Internet reading activities. *Learning and Individual Differences*, 61, 31-39. <https://doi.org/10.1016/j.lindif.2017.11.006>

- *Salmerón, L., Vargas, C., Delgado, P., & Altamura, L. (2020). [Unpublished raw data]. ERI Lectura, University of Valencia.
- Salmerón, L., Vargas, C., Delgado, P., & Baron, N. (2023). Relation between digital tool practices in the language arts classroom and reading comprehension scores. *Reading and Writing*, 36, 175–194. <https://doi.org/10.1007/s11145-022-10295-1>
- *Smith, M. C., & Smith, T. J. (2010). Adults' uses of computer technology: associations with literacy tasks. *Journal of Educational Computing Research*, 42(4), 407–422. <https://doi.org/10.2190/EC.42.4.c>
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328(5977), 450–452. <https://doi.org/10.1126/science.1182597>
- Stadtler, M., Scharrer, L., Brummernhenrich, B., & Bromme, R. (2013). Dealing with uncertainty: Readers' memory for and use of conflicting information from science texts as function of presentation format and source expertise. *Cognition and Instruction*, 31(2), 130–150. <https://doi.org/10.1080/07370008.2013.769996>
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407. <https://doi.org/10.1598/RRQ.21.4.1>
- Stothard, S. E., Hulme, C., Clarke, P., Barmby, P., & Snowling, M. J. (2010). *York assessment of reading for comprehension. Secondary test*. London, UK: GL Assessment.
- SweSAT-R. (2006). *Swedish Scholastic Aptitude Test Reviewed*. Stockholm, Sweden.
- Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, 40(6), 604–34. <https://doi.org/10.3102/1076998615606099>
- *Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M. K., Tolvanen, A., & Poikkeus, A. M. (2019). Leisure reading (but not any kind) and reading comprehension support each other—A longitudinal study across grades 1 and 9. *Child Development*, 91(3), 876–900. <https://doi.org/10.1111/cdev.13241>
- Twenge, J. M., Martin, G. N., & Spitzberg, B. H. (2019). Trends in U.S. Adolescents' media use, 1976–2016: The rise of digital media, the decline of TV, and the (near) demise of print. *Psychology of Popular Media Culture*, 8(4), 329–345. <https://doi.org/10.1037/ppm0000203>

- van Bergen, E., Vasalampi, K., & Torppa, M. (2020). How are practice and performance related? Development of reading from age 5 to 15. *Reading Research Quarterly*, 56(3), 415–434. <https://doi.org/10.1002/rrq.309>
- Van den Noortgate, W., López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, 45, 576–594. <https://doi.org/10.3758/s13428-012-0261-6>
- *Vergara, M., Delgado, P., Salmerón, L., Altamura, L. (2020). [Unpublished raw data]. ERI Lectura, University of Valencia.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Viechtbauer, W. (2021). Model checking in meta-analysis. In C. H. Schmid, T. Stijnen, & I. R. White (Eds.), *Handbook of meta-analysis* (pp. 219-254). CRC Press. <https://doi.org/10.1201/9781315119403>
- Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>
- Warrens M. J. (2010). A formal proof of a paradox associated with Cohen’s kappa. *Journal of Classification*, 27(3), 322–332. <https://doi.org/10.1007/s00357-010-9060-x>
- Wiederholt, J. L., & Blalock, G. (2000). *Gray Silent Reading Tests GSRT*. Pro-Ed.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test 3 (WRAT3)*. Wilmington, DE: Wide Range, Inc.
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test 4 (WRAT4)*. Lutz, FL: Psychological Assessment Resources.
- Wolf, M. (2018). Reader, come home: The reading brain in a digital world (p. 54). New York: Harper.
- *Wu, J. Y., & Peng, Y. C. (2017). The modality effect on reading literacy: Perspectives from students’ online reading habits, cognitive and metacognitive strategies, and web navigation skills across regions. *Interactive Learning Environments*, 25(7), 859-876. <https://doi.org/10.1080/10494820.2016.1224251>

Pre-published version

Table 1

Descriptive characteristics of the digital reading habits of the studies included

Study	N of items	Items	Scale	Category
Altamura (2019)	4	Emails, online messages (e.g., WhatsApp), social networks, blogs	5-point Likert scale (Never, Few days per month, Once a week, Few days per week, Every day) - Times per day/week/month	Social-communicative use of reading
Amiama-Espaillet & Mayor-Ruiz (2017)	10	Download files, books, fanfiction, magazines or blogs, video, email, consult, virtual platform, virtual shopping, social networks	5-point Likert scale (from 1: Never to 5: Always) - Frequency adverbs	Linear-informational reading
Champley et al. (2008)	1	Electronic reading resources	5-point Likert scale (1: Never to 5: Often) - Frequency adverbs	Linear-informational reading
Chen & Fang (2013)	4	E-fiction, e-news, blogs and Bulletin Board Systems	4-point scale (from 0: Never to 3: Frequently) - Frequency adverbs	Linear-informational reading
Duncan et al. (2016)	6	Factual website or blog (e.g., Wikipedia), online searching, text messages/e-mails, networking website (e.g., Facebook), computer/console game, and Twitter	5-point Likert scale (1: Didn't read this, 2: 30 min or less, 3: 1 hr, 4: 2 hr, 5: 3 hr or more) – Hours (weekly)	Mixed or unspecified
García (2013)	5	Browsing the internet for fun (e.g. YouTube), read or write emails to friends, write a blog or website, chatting online (e.g., Messenger), participate in online forums, participate in social networks (e.g., Twitter, Facebook)	5-point Likert scale (Almost every day, Once or twice a week, A few times a month, Once or twice a month, to Never) - Times per day/week/month	Social-communicative use of reading
Gumus & Atalmis (2011)	6	Browse the Internet for information about people, things, or ideas, Play games, Use the Internet to collaborate with a group or team, Download software from the Internet (including games), Download music from the Internet, For communication (e.g., Email or “chat rooms”)	5-point Likert scale (Almost every day, Once or twice a week, A few times a month, Once or twice a month, to Never) - Times per day/week/month	Mixed or unspecified

Heesbeen (2021)	1	“How often do you read digital texts after schooltime, for pleasure or personal interests (e.g., blogs, e-mail, or messages in social media)?”	4-point-Likert scale 1 (‘Almost never’) to 4 (‘almost every day’) - Times per day/week/month	Mixed or unspecified
Hsu et al. (2019)	4	Reading e-books, reading on computer, reading on smartphones, texting friends	5-point Likert scale (1: None, 2: <2h, 3: 2-3h, 4: 3-5h, 5: >5h) – Hours (daily)	Linear-informational reading
Johnson (2013)	3	Email, play games, and visit websites	4-point Likert scale (Every day or almost every day, Once or twice a week, Once or twice a month, to Never or hardly ever) Times per day/week/month	Mixed or unspecified
Kanniainen et al. (2022)	5	e-books, online newspapers (examples of Finnish online newspapers), websites on different topics (e.g., interests, hobbies, sports), blog postings, and forums (e.g., discussions of games, artists, hobbies)	5-point Likert scale (Hardly ever; 1–2 times per month, 1–2 times per week, Almost every day, Every day) - Times per day/week/month	Linear-informational reading
Kong et al. (2022)	6	Reading emails, reading online news, chatting online, online information searches to learn about a particular topic, participating in online group discussions, and searching for practical information online	5-point Likert scale (Every day, almost every day, once or twice a week, once or twice a month, to never or hardly ever) - Times per day/week/month	Mixed or unspecified
Macedo-Rouet et al. (2019)	2	Communication with family or friends through social networks (Facebook, Twitter ...), looking for information on the Internet about subjects of personal interest (sport, leisure, music, general culture, etc.)	5-point Likert scale (Never or I don’t know what it is, Never or hardly ever, About once a month, Once or twice a week, Almost every day, At least once a day, Several times a day) - Times per day/week/month	Mixed or unspecified
McGeown et al. (2016)	5	Online searching/web browsing, text messages, emails, social networking site (e.g., Facebook) and Twitter	5-point Likert scale (Don’t read this, Less than 1 hour, About 1–2 hours, About 3–4 hours and 5 hours or more) – Hours (weekly)	Social-communicative use of reading

Mežek (2013)	1	“In an average week in term time, how much time do you spend on leisure reading from the Internet sources you mentioned?”	5-point Likert scale (Less than 3.5 hours (30 min a day), 3.5-7 hours, 8-14 hours, 15-21 hours, More than 21 hours) – Hours (daily)	Mixed or unspecified
Norris (2010)	4	Email, chat, surf, video games	4-point Likert scale (Less than once a week, about once a week, several times a week, and every day or almost every day) - Times per day/week/month	Social-communicative use of reading
Park & Weng (2020)	12	PISA 2012 ENTUSE items + Playing online games via social networks (e.g., <Farmville®>, <The Sims Social>), Downloading new apps on a mobile device	5-point Likert scale (Every day, almost every day, once or twice a week, once or twice a month, to never or hardly ever) - Times per day/week/month	Mixed or unspecified
Petko et al. (2017)	10	PISA 2012 ENTUSE items: Playing one-player games; Playing collaborative online games; Using email; <Chatting online> (e.g., <MSN®>); Participating in social networks (e.g., <Facebook>, <MySpace>); Browsing the Internet for fun (such as watching videos, e.g., <YouTube™>); Reading news on the Internet (e.g., current affairs); Obtaining practical information from the Internet (e.g., locations, dates of events); Downloading music, films, games or software from the Internet; Uploading your own created contents for sharing (e.g., music, poetry, videos, computer programs)	5-point Likert scale (Every day, almost every day, once or twice a week, once or twice a month, to never or hardly ever) - Times per day/week/month	Mixed or unspecified
Pfost et al. (2013)	3	E-mails; online encyclopedias (e.g., Wikipedia); and online forums or chats	4-point Likert scale (Almost never or never, several times a month, several times a week, several times a day) - Times per day/week/month	Mixed or unspecified
Salmerón et al. (2011)	4	Browsing the internet for fun (e.g., YouTube), read or write emails to friends, write a blog or website, chatting online (e.g., Messenger), participate in social networks (e.g., Twitter, Facebook)	5-point Likert scale (Almost every day, once or twice a week, a few times a month, once or twice a month, to never) - Times per day/week/month	Social-communicative use of reading

Salmerón et al. (2018)	6	Browsing the internet for fun (e.g., YouTube), read or write emails to friends, write a blog or website, chatting online (e.g., Messenger), participate in online forums, participate in social networks (e.g., Twitter, Facebook)	5-point Likert scale (Almost every day, once or twice a week, a few times a month, once or twice a month, to never) - Times per day/week/month	Social-communicative use of reading
Salmerón et al. (2020)	8	Digital novels, digital newspapers, digital books about subjects of personal interests	5-point Likert scale (Every day, few times a week, once a week, few times a month, never) - Times per day/week/month	Linear-informational reading
Smith & Smith (2010)	3	Send or receive e-mail, find information on the internet, talk in Internet-chat groups	5-point Likert scale (Every day, few times a week, once a week, few times a month, never) - Times per day/week/month	Social-communicative use of reading
Torppa et al. (2019)	4	E-mails, messages and Internet conversations, Facebook, and blogs	5-point Likert scale (from Never to Daily) - Times per day/week/month	Social-communicative use of reading
Vergara et al. (2021)	1	Reading on digital devices	Open: weekly hours	Linear-informational reading
Wu & Peng (2017)	6	E-mail, chatting online, online news, searching for a particular topic, searching for practical information, and using online dictionaries or Wikis	5-point Likert scale (from 1: I don't know what it is, to 5: Several times a day) - Times per day/week/month	Social-communicative use of reading

Table 2*Descriptive characteristics of the studies included in the analysis*

Study	r	N	Extrinsic data		Sociodemographic data			Reading comprehension assessment		
			Published	Data coll. year	Mean age (SD)	% Fem.	Student level	Name of the test	Genre of the text(s)	Type of questions and assessment
Altamura (2019) Study 1	-.266	35	No: Raw data from bachelor thesis	2019	12.11 (0.53)	57.10	Middle school	PROLEC-SE-R (Cuetos et al., 2016)	Expository	Mixed: inferential and literal 10 Multiple-choice questions and 10 open-ended questions
Altamura (2019) Study 2	.055	47		2019	13.42 (0.87)	53.20	Middle school			
Altamura (2019) Study 3	-.098	187		2019	14.16 (0.68)	49.70	High school			
Altamura (2019) Study 4	.027	168		2019	15.03 (0.61)	46.40	High school			
Altamura (2019) Study 5	-.202	119		2019	10.46 (0.553)	42.00	Primary school	PROLEC-R (Cuetos et al., 2014)	Expository and narrative	Inferential 16 open-ended questions
Altamura (2019) Study 6	-.219	85		2019	11.62 (0.38)	44.70	Middle school			
Amiama-Españat & Mayor-Ruiz (2017)	.253	382	Yes: Research paper	NR	15.15 (0.85)	59.00	High school	Adapted: Based on PISA Texts	Descriptive, expository argumentative, exposition and diagram	Mixed: location, integration and evaluation 13 multiple-choice questions, 6 open-ended and 3 double-entry tables
Champley et al. (2008)	.220	96	Yes: Research paper	2004	72.5 (4.19)	73.90	No students	Qualitative Reading Inventory-3 (Leslie & Caldwell, 2001)	Narrative, expository	Mixed: explicit and implicit
Chen & Fang (2013)	.120	352	Yes: Research paper	2011	NR	59.20	Undergraduates	Reading Comprehension Test developed by Chen and Su (2010)	NR	Mixed: literal (e.g., lexical access and parsing) and inferential (e.g., integration, summarisation and analogy) comprehension

Duncan et al. (2016) Study 1	.063	201	Yes: Research paper	2009	12.66 (0.89)	57.80	Middle school	YARC Secondary test (Stothard et al., 2010)	Fiction and non-fiction	Mixed: Vocabulary, literal, or inferential. 12 questions
Duncan et al. (2016) Study 2	-.041	99		2009	15.09 (0.53)	50.50	High school			
García (2013) Study 1	.192	23	No: Doctoral thesis	2013	11.5	47.36	Middle school	TPC (Martínez et al., 2008)	Expository	Mixed: explicit ideas (literal) comprehension, inference making between different paragraphs ideas, and between texts ideas and readers' prior knowledge 20 multiple-choice questions
García (2013) Study 2	.054	142		2012	11.6	51.00	Middle school			
Gumus & Atalmis (2011)	.425	4,942	Yes: Research paper	2006	15.9	46.30	High school	2006 PISA Reading assessment (OECD, 2006)	Narrative, expository, descriptive, argumentative/persuasive, instructive	Mixed: retrieve information, interpret information, reflect on text content, reflect on the structure of the text Multiple-choice, short closed-constructed response, and open- (extended) constructed response
Heesbeen (2021)	.060	108	No: Doctoral thesis	2021	11.1 (0.7)	47.20	NA	Cito Reading Comprehension test (CITO; Centraal Instituut voor Toetsontwikkeling; Institute for Test Development Netherlands, 2016).	Fictional story and formal text	Mixed: fill in the blank, prediction of content, text content, searching in sources, and summarizing Multiple choice questions
Hsu et al. (2019)	-.127	50	Yes: Research paper	2019	22.88 (4.71)	52.00	Undergraduates	Gray Silent Reading Test - GSRT (Wiederholt & Blalock, 2000)	Narrative	NR 5 multiple-choice questions
Johnson (2013)	.207	90	Yes: Research paper	NR	10.63 (1.3)	34.50	NA	Adapted from the Wide Range Achievement Test Sentence Comprehension subtest (Wilkinson & Robertson, 2006)	NR	NR Cloze technique
Kannianen et al. (2022)	.090	426	Yes: Research paper	2014	12.34 (0.33)	47.00	Middle school	ORCA (Caccia et al., 2019) – Online assessment	Expository	Mixed: locating information, evaluating information, synthesizing information, and communicating information. Process and open tasks

Kong et al. (2022)	.130	14,087	Yes: Research paper	2018	15	49.10	High school	Digital reading comprehension 2018 PISA Assessment (OECD, 2019)	Descriptive, narrative, expository, argumentative, instructive, transaction texts	Mixed: scanning and locating, literal comprehension, inferential comprehension, assessing quality and credibility, reflecting on content and form, searching for and selecting relevant text, multiple-text inferential comprehension, corroborating/handling conflict Multiple choice and open constructed-response
Macedo-Rouet et al. (2019)	-.005	146	Yes: Research paper	2018	14.73 (0.53)	53.40	High school	Protocole Emilie (Duchêne May Carle, 2010)	Narrative	NR True-false
McGeown et al. (2016)	-.082	784	Yes: Research paper	2012	9.91 (0.88)	52.70	Primary school	Macmillan Test Unit, Group Reading Test II (2000)	NR	NR Cloze technique
Mežek (2013)	.075	31	Yes: Research paper	2009	NA	85.30	Undergraduates	SweSAT-R (2006)	NR	NR 4 multiple-choice questions
Norris (2010)	.059	515	No: Doctoral thesis	2004	11.25 (6.34)	54.00	NA	Peabody Individual Achievement Test (PIAT Comprehension) (Dunn & Markwardt, 1970)	NR	NR
Park & Weng (2020)	-.016	168,098	Yes: Research paper	2015	15.78	NR	High school	2015 PISA Reading assessment (OECD, 2017)	Narrative, expository, descriptive, argumentative/persuasive, instructive	Mixed: retrieve information, interpret information, reflect on text content, reflect on the structure of the text.
Petko et al. (2017)	.078	222,811	Yes: Research paper	2012	15.77 (0.048)	NR	High school	2012 PISA Reading assessment (OECD, 2013)		Multiple-choice, short closed-constructed response, and open- (extended) constructed response.
Pfost et al. (2013)	-.117	1,124	Yes: Research paper	2010	13.4	52.70	Middle school	NA	Narrative and expository	High inferences 26 multiple-choice
Salmerón et al. (2011) Study 1	-.093	74		2011	12.36 (0.538)	51.90	Middle school	WebLEC (Salmerón et al., 2018) and CompLEC (Llorens et al.,	Expository	

Salmerón et al. (2011) Study 2	.141	58		2011	13.49 (0.658)	51.90	Middle school	2011) (we used the CompLEC scores)		
Salmerón et al. (2011) Study 3	-.029	77	No: Unpublished raw data	2011	14.48 (0.749)	44.90	High school			
Salmerón et al. (2011) Study 4	.236	64		2011	15.48 (0.651)	41.50	High school			
Salmerón et al. (2018) Study 1	-.209	138		2015	12.13 (0.542)	41.50	Middle school			Mixed: retrieve information, interpret information, reflect on text content, reflect on the structure of the text
Salmerón et al. (2018) Study 2	-.137	132	Yes: Research paper	2015	13.17 (0.601)	45.90	Middle school			17 multiple-choice questions and 3 open-ended
Salmerón et al. (2018) Study 3	.068	136		2015	14.18 (0.651)	50.00	High school			
Salmerón et al. (2018) Study 4	.007	136		2015	15.13 (0.579)	47.80	High school			
Salmerón et al. (2020) Study 1	.118	191	No: Unpublished raw data	2020	20.95 (3.959)	86.90	Undergraduates	CLOZE (Salmerón et al., 2022)	Narrative	High inference from little context
Salmerón et al. (2020) Study 2	.063	120		2020	22.76 (5.593)	80.70	Undergraduates			Cloze technique (34 items)
Smith & Smith (2010)	.311	17,668	Yes: Research paper	2003	NA	57.10	Mix	NAAL literacy assessment (NAAL; NCES, n.d.a)	Prose literacy. Prose texts can be further broken down as expository, narrative, procedural, or persuasive. Document and quantitative literacy	NR
Torppa et al. (2019)	-.060	1,813	Yes: Research paper	2013	12.5	NA	Middle school	YKÄ (Lerkkanen et al., 2018)	NR	NR
Vergara et al. (2021)	.045	113	No: Unpublished raw data	2021	20.02 (1.88)	92.00	Undergraduates	CLOZE (Salmerón et al., 2022)	Narrative	High inference from little context Cloze technique (34 items)

Wu & Peng (2017)	.122	33,696	Yes: Research paper	2009	15.77	50.78	High school	Printed Reading Assessment (PRA), Digital Reading Assessment (DRA) from 2009 PISA tests (OECD, 2010) (we used the PRA scores)	Narrative, expository, descriptive, argumentative/persuasive, injunctive	Mixed: retrieve information, interpret information, reflect on text content, reflect on the structure of the text Multiple-choice, short closed-constructed response, and open- (extended) constructed response
------------------	------	--------	---------------------	------	-------	-------	-------------	---	--	--

Note. r : effect size, N : number of participants, Data Coll. Year: Data Collection Year, % Fem: Percentage of female participants, NA: Not Applicable, NR: Not reported

Table 3

Results of the mixed-effects meta-regressions for the qualitative moderator variables on the effect sizes obtained from the studies

Variable	<i>k</i>	<i>n</i>	<i>b</i>	SE	95% C.I. [LL, UL]	<i>F</i> (df)	<i>Qe</i> (df)	Pseudo <i>R</i> ²	<i>p</i>
<i>Sociodemographic variables</i>									
Educational stage	36	22				13.522(3, 2.44)	1238.47(32)***	.57	.047
Primary/Elementary ¹	2		-.095	.018	[-.32, .13]				
Middle school ²	14		-.025	.035	[-.10, .06]				
High school ²	14		.085	.026	[.01, .03]				
Undergraduate ¹	6		.070	.03	[-.03, .17]				
<i>Methodological variables</i>									
Extra-large sample size	39	25				2.976(2, 8.72)	2554.687(37)***	.12	.153
< 10.000	34		.029	.029	[-.02, .08]				
>10.000	5		.127	.056	[-.02, .28]				
Sampling method	35	22				.0851(1, 17.52)	2573.99(33)***	.00	.774
Non-probability	26		.06	.037	[-.01, .13]				

Probability	9		.08	.041	[.003, .15]		
<i>Digital reading habits instrument characteristics</i>							
Temporal ranges	38	34			1.919 (4, 5.43)	1540.57 (34)***	.27 .204
Non-specified	5		.11	.07	[-.09, .30]		
Daily	4		.21	.07	[-.06, .48]		
Weekly habits	5		.00	.04	[-.11, .12]		
Mixed	24		.03	.03	[-.05, .07]		
Temporal classification	38	24			.1791(2, 5.04)	2578.79(35)***	.00 .841
Hours	5		.03	.064	[-.20, .26]		
Frequency adverbs	5		.09	.074	[-.12, .31]		
Times per day/week/month	28		.05	.031	[-.02, .12]		
Number of answer alternatives	38	24			.1305 (2, 1.66)	2577.87(35)***	.00 .886
4-point Likert-scale	4		.02	.058	[-.16, .21]		
5-point Likert-scale	32		.05	.031	[-0.006, .12]		
7-point Likert-scale	2		.09	.107	[.03, .26]		
Type of digital reading habit	39	25			1.499(2, 13.35)	1740.73(36)***	.02 .258

Social-communicative	21		.03	.048	[-.08, .14]			
Linear-informative	8		.12	.037	[.02, .21]			
Mixed	10		.03	.033	[-.04, .11]			
<i>Reading comprehension test characteristics</i>								
Media	32	18				.0684 (1, 8.84)	1373.97(30)***	.00 .799
Printed	25		.07	.041	[-.02, .16]			
Digital-based	7		.05	.031	[-.03, .14]			
Text genre	34	20				.3701 (2, 5.06)	2545.36(31)***	.00 .708
Narrative	5		.02	.042	[-.12, .16]			
Expository	15		.06	.04	[-.04, .16]			
Mixed	14		.08	.05	[-.02, .18]			
Assessment method	35	21				.4867 (3, 4.83)	1341.74(31)***	.00 .707
Multiple-choice questions (MCQ)	8		.01	.049	[-.10, .14]			
Open-ended	6		.05	.011	[-.27, .39]			
CLOZE type	4		.01	.060	[-.25, .28]			
Mixed	17		.09	.034	[.008, .17]			

Reading comprehension dimension	31	17				5.169 (1, 3.35)	1241.34(29)***	.44	.098
Inferential	6		-.05	.058	[-.24, .14]				
Literal	0								
Mixed	25		.08	.022	[.03, .13]				
<i>Extrinsic variables</i>									
Published status	39	25				.4312(1, 10.38)	2588.62(37)***	.00	.528
Non-published	17		.03		[-.05, .12]				
Published	22		.06		[-.005, .13]				

Note. k : number of effect sizes. n : number of studies (clusters). r : mean effect size. F : statistical test for testing the significance of the moderator variable. Q_E : statistic for testing the model misspecification. $pseudo R^2$: proportion of variance explained by the moderator (τ^2 between - τ^2 within) ^{1,2} Significant paired comparisons. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Results of the mixed-effects meta-regressions for the continuous moderator variables on the effect sizes obtained from the studies

Variable	<i>k</i>	<i>n</i>	<i>b</i>	<i>SE</i>	95% <i>C.I.</i>	<i>F(df)</i>	<i>Q_{E(df)}</i>	<i>R</i> ²	<i>p</i>
<i>Sociodemographic variables</i>									
Mean age of student sample	34	21	.012	.001	[-.01,.03]	2.32(1,6.44)	1293.22(32)***	.38	.175
Mean age of student sample from Primary and Secondary school	30	18	.028	.011	[.002, .05]	6.04(1, 9.12)	1289.91(28)***	.36	.036
Percentage of female	35	21	.003	.002	[-.002, .01]	2.53(1, 4.35)	533.53(33)***	.00	.181
Percentage of male	34	21	-.003	.002	[-.01, .002]	2.54(1, 4.30)	538.48(32)***	.00	.181
<i>Methodological variables</i>									
Sample size	39	25	.00	.00	[-.00, .00]	-. ²	2298.81(37)***	.00	.965
Reading comprehension test reliability	35	21	.326	.265	[-.29, .95]	1.51(1, 7.35)	1020.92(33)***	.31	.257
Digital reading questionnaire reliability	30	17	-.012	.140	[-.36, .34]	.007(1, 5.57)	240.41(28)***	.13	.935
Number of items (DRH)	39	25	.00	.01	[-.02, .02]	0.001(1, 6.21)	874.86(37)***	.00	.969
<i>Extrinsic variables</i>									
Collection data year	38	24	-.01	.006	[-.02, .003]	2.978(1, 8.43)	826.69(36)***	.34	.121

Note. k : number of effect sizes. n : number of studies. b : unstandardized regression coefficient. SE : Standard error. 95% $C.I.$: [LL, UL]. F : statistical test for testing the significance of the moderator variable. Q_E : statistic for testing the model misspecification. $pseudo R^2$: proportion of variance explained by the moderator (τ^2 between - τ^2 within). ² Not statistical power enough to load the calculus. $*p < .05$. $**p < .01$. $***p < .001$.

Table 5

Results of the mixed-effects meta-regression analysis for the average effect sizes obtained from the studies

Analysis	r	95% C.I. [LL, UL]	p	95% P.I. [LL, UL]	τ^2 between	τ^2 within
Main Analysis	0.073	[.012, .013]	.019	[-.21, .35]	.014	.004
Infl. Cases Removed ¹	0.055	[.003, .107]	.039	[-.18, .29]	.009	.004

Note: r : mean effect size, C.I: Confidence Interval, P.I: Prediction interval, τ^2 between: between-study variance, τ^2 within: within-study variance.

¹Removed as outliers: Gumus & Atalmis (2011).