




Article

Medium Matters? Comprehension and Lexical Processing in Digital and Printed Narrative Texts in Good and Poor Comprehenders

Elisabetta Lombardo ^{1,*}, Ambra Fastelli ^{1,2}, Sara Gaudio ¹ and Paola Bonifacci ¹

¹ Department of Psychology, University of Bologna, Viale Berti Pichat n.5, 40127 Bologna, Italy; ambrafastelli.psicologa@gmail.com (A.F.); sara.gaudio2@unibo.it (S.G.); paola.bonifacci@unibo.it (P.B.)

² National Institute for the Evaluation of the Education and Training System—INVALSI, Via Ippolito Nievo, 35, 00153 Roma, Italy

* Correspondence: elisabetta.lombardo@unibo.it

Abstract

The present study examined differences in reading comprehension performance between good and poor comprehenders, across paper-based and computer-based formats. The sample consisted of 197 students ($M_{age} = 10.9$, $SD_{age} = 1.22$), categorized into three groups based on their reading comprehension proficiency: good ($n = 73$), average ($n = 90$), and poor ($n = 33$). Using a pseudo-randomized within-subjects design, participants read two texts and completed both a cloze task and a proofreading task in paper and digital formats. Results showed that poor comprehenders consistently performed worse on both tasks; however, group performances were not influenced by the modality. Both tasks required more time in the digital modality and were associated with greater calibration bias. In the proof-reading task, nouns and adjectives were more difficult to retrieve than verbs and function words, whereas in the cloze task, function words were the easiest to supply. The discussion emphasizes the need to account for task type and linguistic complexity when evaluating comprehension. Importantly, the lack of interaction between reading proficiency and modality suggests that digital assessments are comparably effective and reliable across different levels of reading ability.

Keywords: computer-based tasks; reading comprehension; poor comprehenders; primary school; metacognitive variables



check for updates

Academic Editors: Nance S. Wilson and Brittany Adams

Received: 20 May 2025

Revised: 28 July 2025

Accepted: 29 July 2025

Published: 3 August 2025

Citation: Lombardo, E., Fastelli, A., Gaudio, S., & Bonifacci, P. (2025). Medium Matters? Comprehension and Lexical Processing in Digital and Printed Narrative Texts in Good and Poor Comprehenders. *Education Sciences*, 15(8), 989. <https://doi.org/10.3390/educsci15080989>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ubiquitous use of digital devices in the educational contexts and assessment procedures has highlighted the necessity to investigate the role of the medium in fostering reading comprehension processes, whether paper-based or computer-based. While digital tools open up to new ways to interact with the reading material (Mangen, 2016), many studies investigated their impact on reading comprehension outcomes in younger and older students.

As a consequence of this, some evidence highlighted the disadvantage of screen-based texts and comprehension tasks compared to paper ones (screen inferiority—Ackerman & Lauterman, 2012). As observed by Clinton (2019), the effect of the medium is generally addressed in two main dimensions: performances and process. Performance is accounted for by comprehension outcomes and accuracy during the comprehension task, while process is investigated through reading times and metacognitive judgements.

A foundational study by [Delgado et al. \(2018\)](#) provides a meta-analysis that examines the implications of reading media on comprehension, finding that comprehension scores can be significantly impacted by the medium in which a text is presented, with an advantage of the paper modality. More recent studies have shown how digital comprehension may be impacted by different metacognitive engagements in the comprehension process and monitoring ([Ackerman & Goldsmith, 2011](#); [Florit et al., 2022](#)) compared to paper, leading to overconfidence in the digital modality and higher discrepancy between the perceived success of the learning experience and the actual performance obtained ([Ronconi et al., 2022](#)). Moreover, under self-paced modalities, students tend to spend less time reading digitally compared to paper, even when performing less accurately in comprehension. [Delgado and Salmerón \(2021\)](#) investigated how the reading medium affects attention and comprehension under time pressure, indicating that on-screen readers tend to exhibit higher instances of mind-wandering, which hinders comprehension performance. Moreover, since digital reading is often associated with shallower processing for quick rewards and relatively simple reading tasks, the digital context might be associated with leisure more than deep reading and activate these same strategies during more complex comprehension tasks ([Salmerón et al., 2022](#); [Annisette & Lafreniere, 2017](#)—shallowing hypothesis).

Nonetheless, these results are far from being conclusive. Many factors might play a role in the expression of screen inferiority ([Peras et al., 2023](#)). On the one hand, contextual factors, such as the learning environment (formal versus informal, individual versus group learning), content of the text (expository versus narrative), and the requests of the task (deeper or superficial information processing) concur with the expression of the effects of the medium on comprehension performance ([Clinton, 2019](#); [Singer & Alexander, 2017](#); [Lenhard et al., 2017](#)). The recent systematic review of [Peras et al. \(2023\)](#) suggests that prior comprehension skills and environmental factors significantly affect comprehension when engaging with digital media. They highlight that students who struggle with reading comprehension in traditional formats may face heightened difficulties in digital contexts, advocating for tailored instructional strategies. Research is needed to understand how the impact of the medium on reading comprehension performances might differ among good, average, and poor comprehenders, identified by the use of standardized instruments of evaluation.

1.1. Cognitive and Metacognitive Processes in Impaired Reading Comprehension

Reading comprehension is a multidimensional process that requires different cognitive skills, at different levels. On the one hand, reading comprehension is the result of the reader's ability to decode written words, recall their meaning, assess their relationships, and create a mental and integrated representation of the text ([Kintsch & Kintsch, 2005](#); [Kintsch & van Dijk, 1978](#)). According to the Simple View of Reading (SVR) model, reading comprehension is the product of efficient decoding skills and language comprehension skills ([Gough & Tunmer, 1986](#); [Hoover & Gough, 1990](#)). However, in order to achieve a deep understanding of a text, the reader must rely on prior knowledge about the world and the topic at hand, in addition to the way texts are constructed to convey meaning.

As the reading-comprehension process is complex and multifaced, many are the reasons why it could be impaired in young readers. Specific difficulties in comprehension encompass the ability to decode written text correctly while struggling in accessing vocabulary information, using metacognitive strategies, and drawing inferences—all critical elements for successful reading comprehension ([Nation & Snowling, 1998](#); [Yuill & Oakhill, 1991](#)). These difficulties seem to be unrelated to the type of text (narrative or descriptive) or the type of comprehension level being tested, local or global ([Bonifacci & Tobia, 2016](#)).

Poor comprehenders may have associated weaknesses in other cognitive processes. On the one hand, studies showed that poor comprehenders perform lower than their peers in complex verbal working memory tasks, particularly in tasks that require both storage and processing of information (Georgiou & Das, 2014; Carretti et al., 2009). More specifically, poor comprehenders perform similarly to their peers on short-term memory tasks (e.g., digit span) but struggle in tasks that require the manipulation of information rather than mere recall (Cragg & Nation, 2006). This distinction highlights the fact that while their short-term memory capacity appears adequate, the quality of their cognitive processing during working memory tasks is compromised. This is also in line with another study that links the difficulties of these students in inferring meanings and retaining newly acquired vocabulary to inadequate morphological and syntactic awareness (Tong et al., 2013).

Alongside cognitive processes, metacognition also plays an important role during text comprehension. Effective reading comprehension requires the ability to correctly assess one's level of understanding while reading, compared to previously self-imposed or externally stated reading goals. Research has shown that students who engage in more accurate metacognitive judgments about their understanding tend to achieve higher levels of comprehension (Mirandola et al., 2018). This underscores the importance of calibration between a reader's self-assessment and actual performance, as it allows learners to identify difficulties encountered during reading and to employ appropriate strategies for resolution. Evidence suggests that deficits in metacognitive awareness can influence these students' comprehension skills, as well as a lack of strategies to monitor their understanding (Elwér et al., 2013; Tobia & Bonifacci, 2020).

Given that the individual factors underlying specific comprehension difficulties appear to be many and linked to students' overall cognitive functioning, research has aimed to explore how new technologies can support the implementation of tools designed to aid in text comprehension in digital contexts, with particular attention to the individualization of the learning experience. As such, the ongoing research on individual differences in reading comprehension must now take into account the effects of the medium on this process.

1.2. Poor Comprehenders and the Use of Digital Devices for Learning Outcomes

The implementation of digital tools could allow better access to individualized learning experiences, but limited evidence has been brought out about differences between good and poor comprehenders when addressed through digital assessments.

While there are many reasons to expect an influence of previous reading comprehension abilities on the expression of screen inferiority (Singer & Alexander, 2017) some studies highlighted the positive implication of digitalized assessment practices. A previous study by Ruffini et al. (2023) found that children with low comprehension and writing skills benefited more from the digital modality than from paper during comprehension tasks, in comparison to high performers that showed the opposite pattern. The findings of this study indicate that the manifestation of the screen inferiority effect may vary depending on individuals' skill levels, thereby warranting further empirical exploration. While these results are promising for the implementation of digital devices to provide better access to learning opportunities, further proof is needed.

On the one hand, the role of computers in fostering learning processes could also be enhanced through specific types of interaction with digital tools, such as adding text-to-speech applications that may enhance comprehension skills in good and poor comprehenders (Bonifacci et al., 2021).

On the other hand, while this prospect is promising, further research is needed to clarify which types of interaction with digital tools contribute to the emergence—or mitigation—of screen inferiority, also taking into account individual differences.

Recent investigations by [Furenes et al. \(2021\)](#) provide quantitatively robust insights that illustrate that children who typically show considerable variability in comprehension abilities exhibit significant disparities in learning from digital versus paper mediums. Their meta-analysis indicates that children generally comprehend narrative structures better when reading from paper, especially those who are categorized as poor comprehenders. This seems to be linked to specific difficulties in integrating information when reading on digital platforms. [Mangen et al. \(2019\)](#) have reported how tactile feedback and visual engagement afforded by paper texts are key to fostering deeper comprehension compared to interactions with the digital device. These cues—such as flipping pages and paper texture—seem to be of help in keeping track of the structure of the text. This would explain why, especially when considering long texts read on electronic devices versus paper, participants are more engaged and capable of recalling events and reconstructing narratives when reading printed materials. Moreover, the impact of contextual factors in expressing differences between groups in digital environments should be taken into account. For instance, as demonstrated by [Ben-Yehudah and Eshet-Alkalai \(2020\)](#), the congruence between study and test mediums significantly influences reading comprehension outcomes showing that poor comprehenders often perform worse in tasks where the reading medium does not align with the testing medium. Altogether, poor comprehenders may face specific challenges when reading from screens, which leads to difficulties in integrating what they read. Nonetheless, the effect of media between good and poor comprehenders is still relatively under-researched. These differences might be marked by distinct challenges and advantages posed by these media. While both groups tend to benefit from the use of printed material, differences in metacognitive and cognitive processes might be expressed diversely across modalities. Moreover, most studies to date have analyzed differences between paper and screen by considering overall comprehension scores, but especially considering poor comprehenders, it would be important to better understand to what extent the task demands in terms of semantic processing ([Nation & Snowling, 2008](#)) may vary in relation to the medium.

1.3. The Present Study

The present study investigated differences between good and poor comprehenders in the performances based on reading comprehension tasks administered in both paper and computer formats. The experimental protocol made use of an online platform and an ad hoc testing battery, specifically developed for the purposes of the study, within a broader research project that examined modality-related differences in literacy tasks.

In particular, the study focused on performance in a cloze task and a proof-reading task, both designed to assess the retrieval and recognition of lexical information from different lexical categories (nouns and adjectives, verbs and functors) previously acquired through the reading of a small narrative text. In this study, nouns and adjectives were grouped into a single category due to their shared referential and descriptive functions, as both contribute to representing entities and their properties in narrative comprehension ([Kintsch & Kintsch, 2005](#)). Verbs were considered separately because they encode actions, events, and states, involving different syntactic and semantic processing. Function words (such as articles, conjunctions, and prepositions) formed a third category, given their grammatical role and low semantic content. This classification reflects evidence that content words, verbs, and function words are processed and acquired differently ([Bird et al., 2001](#)), and it enables a more refined analysis of lexical retrieval in tasks like cloze completion, where lexical and grammatical demands vary depending on reader ability.

Additionally, the study examined individual differences in calibration bias—that is, the degree of correspondence between participants' judgments of their own performance

and their actual task efficacy—as well as differences in task completion times, comparing children identified as good and poor comprehenders of a primary and secondary Italian school.

Therefore, the present study aimed to investigate the following research questions:

RQ1. “Do good and poor comprehenders score differently when recalling lexical information in literal comprehension task versus comprehension monitoring task, presented both on digital and on paper?”

As stated by the research (Tobia & Bonifacci, 2020), poor comprehenders show impaired metacognitive and verbal skills compared to good comprehenders. Conversely, based on digital tasks, all students show poorer performances than on paper-based ones (Delgado et al., 2018). For these reasons, we expected an interaction between the group and the modality when assessing performances on the cloze task and proof-reading task. Moreover, we expected differences between tasks depending on the cognitive demands, with higher scores in the cloze task compared to the proof-reading, and an effect of the lexical complexity on students’ retrievals.

RQ2. “Do good and poor comprehenders spend different time on task when recalling lexical information in literal comprehension task versus comprehension monitoring task, presented both on digital and on paper?”

The scientific literature on differences between reading media has consistently considered reading time as a key indicator of the reading process and efficiency (Clinton, 2019; Sidi et al., 2017). However, in educational contexts, a particularly relevant metric is the time spent on a task, as it reflects both cognitive effort and learning efficiency—especially when comparing groups with varying levels of reading comprehension. Therefore, the present study aimed to examine the impact of the reading medium on the time spent on the task.

RQ3. “Do good and poor comprehenders self-evaluate differently when recalling lexical information in literal comprehension task versus comprehension monitoring task, presented both on digital and on paper?”

As emerged in the recent scientific literature (Ackerman & Goldsmith, 2011), we expected that all students would show higher discrepancy between metacognitive judgements and actual performances when assessing their success in the digital modality compared to the paper one (calibration bias). We also expected an interaction with the task type, with greater discrepancies emerging during the proof-reading task compared to the cloze task, due to the higher cognitive demands of the former.

2. Materials and Methods

Sample and design: For the present study, a total sample of 197 students was involved ($M_{age} = 10.9$, $SD_{age} = 1.22$), from the 4th and 5th grades of a primary Italian school and from the 1st and 2nd grades of a lower secondary Italian school. The participation was voluntary, and informed consent was signed by parents and tutors, as well as an ad hoc questionnaire about socio-demographic information of the students’ families. All participants demonstrated a non-verbal IQ above 70, as assessed based on the Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 2004; Italian adaptation: Bonifacci & Nori, 2016). Based on their reading comprehension abilities, students were categorized into three groups: good comprehenders ($n = 73$), average comprehenders ($n = 90$), and poor comprehenders ($n = 33$). Groups were determined according to normative scores on the MT-3 Clinica standardized test (Cornoldi et al., 2016), which evaluates reading comprehension through the administration of a narrative text, differentiated between grades, followed by 12 multiple-choice comprehension questions. Students scoring below the 10th percentile were classified as poor comprehenders, those scoring above the 70th percentile were designated as good comprehenders, and students with intermediate scores were

categorized as average comprehenders (Table 1). Since groups were identified using this test, the group variable will be here reported as MT. The three groups did not differ by age, $F(2,193) = 0.843, p = 0.432$, or for gender ($\chi^2(2) = 2.75, p = 0.253$).

Table 1. Distribution of good, average, and poor comprehenders among the four grades.

| Grade | Good Comprehenders | Average Comprehenders | Poor Comprehenders | Total |
|-------|--------------------|-----------------------|--------------------|-------|
| 4 | 13 | 19 | 11 | 43 |
| 5 | 22 | 20 | 5 | 47 |
| 1 | 21 | 27 | 12 | 60 |
| 2 | 18 | 24 | 5 | 47 |

Grades 4 and 5 are from a primary Italian school, while 1 and 2 are from a lower Italian secondary school.

Procedure: The research was carried out in the school context, as part of a broader project on the effects of the medium on reading comprehension and metacognitive processes. The experimental sessions took place both inside the classrooms and in separate rooms (libraries and ICT classrooms) that were made available from the school. During the first group session, participants completed baseline measures and the standardized test for comprehension. The following two individual sessions involved an ad hoc comprehension battery that involved, together with other measures, the literal comprehension task (*cloze-task*) and the detection of inconsistencies task (proof-reading task), presented on both computer-based and paper-based formats. The test administration followed a counter-balanced order, through a pseudo-randomized within-participants design. Nonetheless, as part of a broader project aimed at validating the online platform, the two tasks were presented in a fixed order. This choice was made to ensure consistency in testing conditions and to evaluate the overall functionality of the system. For the purposes of this study, results regarding the lexical retaining during the literal comprehension task and the detection of inconsistencies task will be addressed. Time measures and calibration assessments were also gained through individual sessions. The study was conducted in accordance with the Declaration of Helsinki and the University of Bologna Bioethical Committee approved the project (Prot. 0391784, 10 December 2024).

Materials: The digital protocol was delivered through the online platform VALGO (Società Cooperativa Anastasis—Bologna). The platform was accessible only on computers, using a personal password. Both the paper-based and computer-based protocol involved the same tasks. The comprehension tasks were based on two narrative texts of original content (text A and B), adapted by grade-level in terms of the word count and the overall readability of the text (Table 2). Readability was assessed based on the READIT-base index (DyLan lab—Dinamiche del Linguaggio—TextTools v2.1.9 [Computational Linguistic Institute, “Antonio Zampolli” (ILC) and Italian national research council (CNR)]). Within the two individual experimental sessions, students completed their task after reading one of the two texts appropriate for their grade, without time limits, presented through a pseudo-randomized order between modalities. The text varied from a minimum of 240 words for the 4th grade of primary school to a maximum of 320 words for the 2nd grade of secondary school.

Literal comprehension: cloze task: This task aimed to assess children’s local comprehension abilities. The original text was summarized to reduce cognitive load and participant fatigue during the cloze task. The summary included 12 empty spaces. Each of them belonged to one of the three lexical categories, (2) nouns and (2) adjectives, (4) verbs, and (4) functors. Children had to recall lexical information from the text and complete each empty space with a single word. They could not look back to the text; therefore, they had to retrieve the lexical information through contextual cues. Each answer was assessed with a dichotomous scoring (0 = wrong answer; 1 = correct answer), with a maximum score of

12. Synonyms were assessed as correct answers, as well as verbs in different conjugations. The Chronbach alpha reliability of the task was 0.66.

Table 2. Experimental design.

| Participant | Session | Modality | Text |
|-------------|---------|----------|--------|
| Student 1 | 1 | Digital | Text A |
| | 2 | Paper | Text B |
| Student 2 | 1 | Digital | Text B |
| | 2 | Paper | Text A |
| Student 3 | 1 | Paper | Text A |
| | 2 | Digital | Text B |
| Student 4 | 1 | Paper | Text B |
| | 2 | Digital | Text A |

Pseudo-randomized experimental design: each student was randomly assigned to a computer- or paper-based modality, and the text content was alternated between sessions.

Inconsistencies detection: proof-reading task: This task aimed to assess the reading-comprehension-monitoring abilities of the students. Children had to find 24 errors (12 orthographic and 12 semantic) hidden in the full text, and they had to correct them with the word that was substituted or with a correct synonym. For the purposes of this study, only the scores obtained in the proof-reading of the semantic errors were considered. Each of them belonged to one of the three lexical categories (2 nouns and 2 adjectives, 4 verbs, and 4 functors). The semantic inconsistencies were implemented by substituting words of the same lexical category of the target but of a semantically distant meaning (e.g., “cats” for “children” or “sweep” for “tell”). For each correct answer a dichotomous scoring was used (0 = not identified target, 1 = correctly identified and correct target), with a maximum score of 12. If the target was wrongly identified or wrongly corrected, the identification was assessed with a 0 score. The Chronbach alpha reliability of the task was 0.72. For the purposes of this study, only the scores obtained in the proof-reading of the semantic inconsistencies were considered.

Calibration measures: To assess the calibration bias, at the end of each comprehension task, a visual aid was used to measure children’s self-evaluation, answering the question “How do you think you did during the task?”. The image featured a horizontal gradient scale resembling a thermometer, ranging from red (on the left) to green (on the right), with facial icons to reinforce the emotional tone—sad face on the left end and happy face on the right end. The red end signifies low self-efficacy (they felt they did poorly), and the green end signifies high self-efficacy (they felt they did well). The distance between the red and green end were considered the extremities of a continuum from 1 to 100. Students were asked to place a marker or bar along the scale by clicking on it (in the computer-based modality) or using the pen (in the paper-based modality) to indicate how successful they felt they were. This method allows for a quick, intuitive, and non-verbal expression of perceived performance or confidence. The distance from the red end to the bar placed by the student was converted to a percentage measure of self-efficacy, and the calibration bias was measured as the results of the subtraction between the measured efficacy during the task and the self-efficacy reported by the student (Glenberg et al., 1987).

Data Analyses: To test our hypotheses a linear mixed model (LMM) analysis was carried out using *jamovi* Software Version 2.6 (The Jamovi Project, 2024) with separated models for the three measures: score, time spent on task, and calibration bias. For each model, the random effects were evaluated controlling for the correlation between measures and the explained variance of the model. For significant interactions, the simple effect analysis was carried out, based on expected moderating effects. Post-hoc analyses were carried out with Bonferroni correction applied to control for Type I error inflation. For

assessing statistical significance, a p -value inferior to 0.05 ($p < 0.05$) and a CI of 95% were considered.

3. Results

3.1. RQ1. “Do Good and Poor Comprehenders Score Differently When Recalling Lexical Information in Literal Comprehension Task Versus Comprehension Monitoring Task, Presented Both Digitally and on Paper?”

To test our hypothesis, we performed a linear mixed model with the modality, lexical category, and task as within factors, the comprehension ability (MT) as the between factor, and the score in each category as the dependent variable (Table 3). To account for developmental differences, age was included in the analysis as a covariate.

Table 3. ANOVA Omnibus for reading comprehension scores.

| | F | df | $df(res)$ | p |
|---|----------|------|------------|----------|
| Modality | 0.41912 | 1 | 191.00052 | 0.51816 |
| Task | 62.86688 | 1 | 191.00054 | <0.00001 |
| MT | 28.04498 | 2 | 191.00078 | <0.00001 |
| Lexical Category | 8.28587 | 2 | 1719.00007 | 0.00026 |
| Age | 19.19314 | 1 | 191.00623 | 0.00002 |
| Modality \times Task | 0.81332 | 1 | 1719.00007 | 0.36727 |
| Modality \times MT | 0.17723 | 2 | 191.00020 | 0.83773 |
| Task \times MT | 1.14525 | 2 | 191.00020 | 0.32032 |
| Modality \times Lexical Category | 0.76595 | 2 | 1719.00007 | 0.46505 |
| Task \times Lexical Category | 61.53302 | 2 | 1719.00007 | <0.00001 |
| MT \times Lexical Category | 1.46859 | 4 | 1719.00007 | 0.20924 |
| Lexical Category \times Age | 2.35155 | 2 | 1719.00007 | 0.09553 |
| MT \times Age | 7.69817 | 2 | 191.00078 | 0.00061 |
| Task \times Age | 7.34085 | 1 | 191.00053 | 0.00735 |
| Task \times Lexical Category \times Age | 12.54447 | 2 | 1719.00007 | <0.00001 |
| Modality \times Lexical Category \times Age | 3.64441 | 2 | 1719.00007 | 0.02634 |

In contrast with the screen inferiority, the effect of the modality was not significant [$F(1,194) = 0.419$, $p > 0.05$]. Results showed a main effect of the comprehension ability [$F(193,2) = 28.044$, $p < 0.001$], as well as a main effect of the lexical category [$F(2,1719) = 8.28$, $p < 0.001$] and of the task [$F(1,191) = 62.86$, $p < 0.001$], with a significant interaction between the task and lexical category, $F(2,1719) = 61.53$, $p < 0.001$, (Figure 1). The main effect of age emerged as well, $F(1,193) = 6.25$, $p = 0.013$, hinting to developmental differences that will be further analyzed.

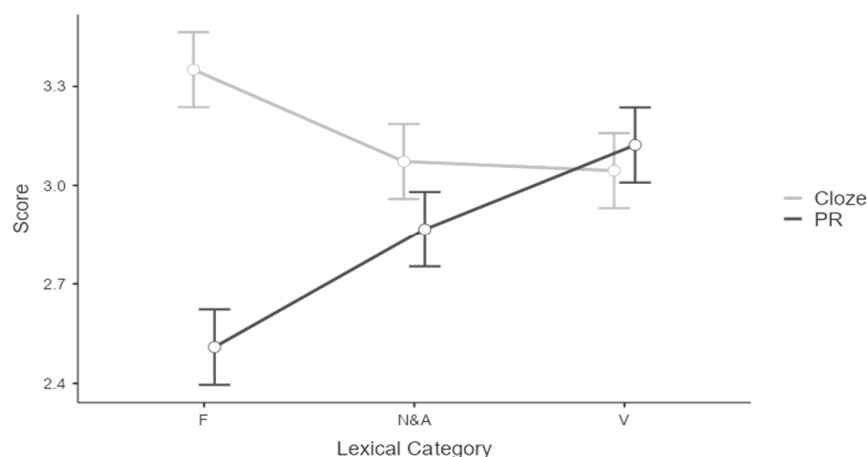


Figure 1. Interaction plot between task and lexical category.

3.1.1. Simple Effect Analysis for Lexical Categories

To examine the influence of the lexical categories within each task, parameter estimates were analyzed. For the cloze task, performance on nouns and adjectives (N&A) was significantly lower than on function words (F) [Estimate = -0.28 , $SE = 0.06$, $t(1719) = -4.7$, $p < 0.001$]. Similarly, performance on verbs (V) was also significantly lower than on function words [Estimate = -0.29 , $SE = 0.06$, $t(1745) = -4.85$, $p < 0.001$].

In contrast, in the proof-reading (PR) task, both nouns and adjectives [Estimate = 0.36 , $SE = 0.06$, $t(1719) = 6.108$, $p < 0.001$] and verbs [Estimate = 0.62 , $SE = 0.06$, $t(1719) = 10.39$, $p < 0.001$] outperformed function words. To further investigate the differences between tasks, the post-hoc analysis adjusted with Bonferroni for the Type I error was carried out on the task variable. Results showed higher scores obtained in the cloze task compared to the proof-reading task, with a mean difference of 0.32 points [$SE = 0.04$, $t(191) = 7.92$, $p < 0.001$].

3.1.2. Differences Between Good Average, and Poor Comprehenders

As emerged in the post-hoc analysis, the experimental protocol was able to show differences among the three groups, independently from the type of format in which the test was presented. More specifically, poor comprehenders were 0.615 points lower than average comprehenders, $t(193) = -5.66$, $p < 0.001$, and 0.879 points lower than good comprehenders, $t(193) = -7.88$, $p < 0.001$. Average comprehenders performed 0.26 points lower than good comprehenders, $t(193) = -3.165$, $p = 0.005$.

Moreover, an interaction with age emerged. The simple effect analysis with age as a moderator showed that the effect of the group is significant, particularly at lower and average levels of age. More specifically, for younger students ($-1SD$ below the mean age), both average comprehenders [Estimates = 0.97 , $SE = 0.14$, 95% CI [0.70, 1.25], $t(191) = 6.97$, $p < 0.001$] and good comprehenders [Estimates = 1.14 , $SE = 0.15$, 95% CI [0.86, 1.43], $t(191) = 7.88$, $p < 0.001$] performed significantly better than poor comprehenders. At the mean age, average comprehenders [Estimates = 0.56 , $SE = 0.11$, 95% CI [0.35, 0.77], $t(191) = 5.26$, $p < 0.001$] and good comprehenders [Estimates = 0.82 , $SE = 0.11$, 95% CI [0.60, 1.03], $t(191) = 7.49$, $p < 0.001$] also significantly outperformed poor comprehenders. Nonetheless, for older students ($+1SD$ above the mean age), only good comprehenders differed significantly from poor comprehenders [Estimates = 0.49 , $SE = 0.17$, 95% CI [0.16, 0.82], $t(191) = 2.95$, $p = 0.004$], while the difference between average and poor comprehenders was not significant [Estimates = 0.15 , $SE = 0.16$, 95% CI [-0.17 , 0.46], $t(191) = 0.92$, $p = 0.358$].

3.1.3. Age Differences Across Tasks

The analysis revealed significant developmental differences in comprehension performance across tasks and lexical categories. Overall, the cloze task was consistently easier than the proof-reading task, particularly among younger participants, with differences decreasing with age. In the computer-based condition, proofreading performance was lower than cloze by -0.43 points for younger, -0.35 for average, and -0.28 for older students (all $p < 0.001$). A similar trend appeared in the paper-based condition, though the task difference was non-significant for older students (Estimate = -0.14 , $p = 0.0764$). Lexical category analysis of the cloze task showed that younger and average-aged students struggled more with nouns and adjectives (up to -0.28) and verbs (up to -0.50) compared to functors, while older students showed no significant difference between verbs and functors ($p = 0.3293$). In contrast, in the proof-reading task, all age groups performed significantly better on nouns and adjectives and verbs than on functors (e.g., mean $-1SD$: $V - F = 0.82$, $p < 0.0001$; mean $+1SD$: $N\&A - F = 0.26$, $p = 0.0045$), indicating that lexical retrieval demands differed by task and age.

3.2. RQ2. “Do Good and Poor Comprehenders Spend Different Time on Task When Recalling Lexical Information in Literal Comprehension Task Versus Comprehension Monitoring Task, Presented Both Digitally and on Paper?”

To test our hypothesis, we performed an ANOVA analysis with modality and task as within factors, the group of performance as the between factor, and the time on task as the dependent variable (Table 4). To account for developmental differences, age was included in the analysis as a covariate. Since time was measured at the end of each task, differences between lexical categories were not assessed.

Table 4. ANOVA Omnibus for time on task.

| | <i>F</i> | <i>df</i> | <i>df (res)</i> | <i>p</i> |
|----------------------------|----------|-----------|-----------------|----------|
| Age | 6.56 | 1 | 191.00 | 0.0112 |
| Modality | 3.82 | 1 | 191.00 | 0.0522 |
| Task | 589.63 | 1 | 191.00 | <0.0001 |
| MT | 2.57 | 2 | 191.00 | 0.0795 |
| Modality × MT | 1.44 | 2 | 191.00 | 0.2387 |
| Task × MT | 1.20 | 2 | 191.00 | 0.3043 |
| Age × Task | 3.49 | 1 | 191.00 | 0.0632 |
| Age × Modality | 7.94 | 1 | 191.00 | 0.0053 |
| Age × MT | 1.64 | 2 | 191.00 | 0.1969 |
| Modality × Task | 5.92 | 1 | 1767.00 | 0.0151 |
| Age × Modality × MT | 3.16 | 2 | 191.00 | 0.0447 |
| Age × Task × MT | 0.81 | 2 | 191.00 | 0.4447 |
| Modality × Task × MT | 5.98 | 2 | 1767.00 | 0.0026 |
| Age × Modality × Task | 1.21 | 1 | 1767.00 | 0.2721 |
| Age × Modality × Task × MT | 15.32 | 2 | 1767.00 | <0.0001 |

While no main effect of comprehension ability emerged, $F(2,191) = 2.57$, $p = 0.079$, there was a marginally significant main effect of modality, $F(1,191) = 3.82$, $p = 0.052$, and significant differences between tasks, $F(1,191) = 589.63$, $p < 0.001$ and age, $F(1,191) = 6.56$, $p = 0.011$. The post-hoc analysis with Bonferroni adjustment was carried out on the contrasts between tasks, showing that the cloze task was carried out 5.94 min faster than the proof-reading task [$SE = 0.24$, $t(191) = -21.28$, $p < 0.0001$], highlighting more effort in the latter task. Moreover, an interaction between age and modality, $F(1,191) = 7.94$, $p = 0.0053$, emerged, as well as a two-way interaction between modality and task, $F(1,1767) = 5.92$, $p = 0.015$.

Even though the main effect of comprehension ability did not emerge, a three-way interaction with modality and age, $F(1,191) = 7.94$, $p = 0.005$, a three-way interaction with modality and comprehension ability, $F(2,191) = 3.16$, $p = 0.044$, and four-way interaction with modality, task, and age, $F(2,1767) = 15.31$, $p < 0.0001$, all emerged. While these results are significant, their interpretation should be guided by caution, since both comprehension ability and modality did not significantly influence the time-on-task variable. Moreover, no significant two-way interaction emerged between task and age. Therefore, the simple effect analyses were carried out on the two-way interaction between age and modality and task and modality.

Simple Effect Analyses on Modality Effect

Firstly, we investigated the moderating role of age on the marginal effect of modality on time spent on task measures, at three levels of the age moderator (mean $- 1 SD$, mean, and mean $+ 1 SD$). Results revealed that while younger students and average students spent more time on paper-based tasks compared to computer-based ones, these differences were non-significant for older students.

More specifically, the analysis revealed that younger students spent 1.01 min less on paper compared to computer-based tasks ($SE = 0.28$, $t(191) = -3.55$, $p = 0.0005$, 95% CI $[-1.57, -0.45]$), while at the mean age, this difference was smaller (Estimate = -0.41 , $SE = 0.21$, $t(191) = -1.95$, $p = 0.0523$, 95% CI $[-0.82, 0.00]$). For older students, there were no significant differences between modalities (Estimate = 0.19 , $SE = 0.31$, $t(191) = 0.60$, $p = 0.5505$, 95% CI $[-0.43, 0.80]$).

Regarding the interaction between task and modality, we investigated the role of task in moderating the role of the medium on the time spent on task measure. The results showed that, while no differences between computer-based and paper-based proof-reading task emerged, $t(226) = -1.18$, $p = 0.24$, the computer-based cloze task were completed 0.56 min faster than the paper-based one [$SE = 0.22$, $t(226) = -2.57$, $p = 0.01$].

3.3. RQ3. “Do Good and Poor Comprehenders Self-Evaluate Differently When Recalling Lexical Information in Literal Comprehension Task Versus Comprehension Monitoring Task, Presented Both Digitally and on Paper?”

To test our hypothesis, we performed an ANOVA analysis with modality and task as within factors, the comprehension ability as the between factor, and the calibration for each task as dependent variable (Table 5). Since the calibration measure was assessed at the end of each task, it was not possible to measure differences between the lexical categories.

Table 5. ANOVA Omnibus for calibration.

| | <i>F</i> | <i>df</i> | <i>df (res)</i> | <i>p</i> |
|----------------------------|----------|-----------|-----------------|----------|
| MT | 1.09 | 2 | 191.00 | 0.3398 |
| Modality | 27.57 | 1 | 191.00 | <0.0001 |
| Task | 37.41 | 1 | 191.00 | <0.0001 |
| Age | 34.09 | 1 | 191.00 | <0.0001 |
| MT × Modality | 1.72 | 2 | 191.00 | 0.1822 |
| MT × Task | 0.21 | 2 | 191.00 | 0.8106 |
| Modality × Task | 127.45 | 1 | 1767.00 | <0.0001 |
| Modality × Age | 0.60 | 1 | 191.00 | 0.4378 |
| MT × Age | 8.45 | 2 | 191.00 | 0.0003 |
| Task × Age | 17.28 | 1 | 191.00 | <0.0001 |
| MT × Modality × Task | 22.38 | 2 | 1767.00 | <0.0001 |
| MT × Modality × Task × Age | 11.76 | 2 | 1767.00 | <0.0001 |

The results showed no significant differences among poor, average, and good comprehenders, $F(2,193) = 1.98$, $p = 0.14$, while the main effect of the task, $F(1,194) = 37.45$, $p < 0.0001$, and the main effect of the modality, $F(1,194) = 26.185$, $p < 0.0001$, emerged. Moreover, a two-way interaction between task and modality, $F(1,1770) = 129.26$, $p < 0.0001$, as well as a two-way interaction between task and age, $F(1,191) = 17.28$, $p < 0.001$, emerged. To further investigate these results, a simple effect analysis was carried out. Firstly, to better understand the effect of medium in the expression of overconfidence, we investigated how the task might moderate the differences between paper and computer in influencing calibration in young students (Figure 2).

The ANOVA for the simple effects of modality showed that while during the cloze task differences between modalities did not emerge [Estimate = -2.34 , $SE = 1.47$, $t(231) = -1.59$, $p = 0.112$], during the proof-reading task students showed less overconfidence on the paper-based task compared to the computer-based one [Estimate = -12.05 , $SE = 1.47$, $t(231) = -8.198$, $p < 0.0001$].

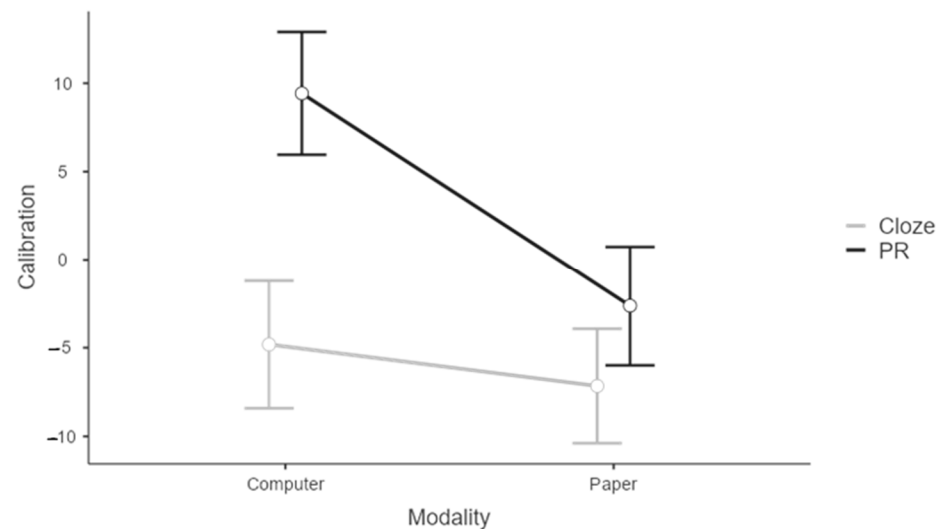


Figure 2. Calibration bias across modalities and tasks.

Then, we decided to look into the effect of the task in the expression of overconfidence at three levels of age as a moderator (mean $- 1 SD$ age, mean age, mean $+ 1 SD$ age). Results showed, at a confidence interval of 95%, that overconfidence is higher in the proof-task compared to the cloze task, both in younger [Estimate = 15.29, $SE = 2.00$, $t(191) = -7.63$, $p < 0.0001$] and averaged-age students [Estimate = 9.07, $SE = 1.48$, $t(191) = -6.11$, $p < 0.0001$]. These differences are no longer significant as the students grow older [Estimate = 2.85, $SE = 2.21$, $t(191) = 1.29$, $p = 0.197$].

While the comprehension abilities level did not show a significant main effect on calibration, results showed several significant interactions with this factor. However, for the purposes of this study, we will focus on the higher-order (i.e., four-way) interaction with modality, task, and age, $F(2,1767) = 11.76$, $p < 0.0001$. Therefore, a simple effect analysis was carried out to understand the moderating role of comprehension abilities, task, and age on the effect of the medium on calibration. The analysis confirmed that, independently from the comprehension abilities, students showed higher overconfidence in the computer-based proof-reading task. Nonetheless, older average comprehenders (mean $+ 1 SD$) showed a different pattern, and both contrasts between computer-based and paper-based cloze task (Estimate = 4.17, $SE = 2.66$, $p > 0.05$) and proof-reading (Estimate = 2.98, $SE = 2.66$, $p > 0.05$) were non-significant.

4. Discussion

The present study investigated how good, average, and poor comprehenders perform in self-paced tasks that measure lexical information recall (cloze task) and comprehension monitoring (proof-reading) across digital and paper modalities. Moreover, time of completion and calibration bias were measured to better understand the underlying metacognitive processes that might influence the expression of differences between modalities and groups. These results contribute to a deeper understanding of how the task type and lexical characteristics of the text interact with individual differences in reading comprehension across modalities (computer versus paper).

4.1. Lexical Recall and Task Performance Between Groups and Modalities (RQ1)

Our findings revealed a robust main effect of the performance group, with good comprehenders consistently outperforming average and poor comprehenders. This confirms previous literature suggesting that reading comprehension is strongly tied to lexical retention and integration processes (Cain & Oakhill, 2006). Interestingly, no significant main effect

of modality was observed, suggesting that the format of task presentation—computer or paper—did not influence the accuracy of lexical retrieval. While this is at odds with the scientific research that brought evidence for screen inferiority and the shallowing hypothesis, these results align with recent findings suggesting that digital media do not inherently impair comprehension, where individual and contextual factors might play a significant role as well (Peras et al., 2023), while on the other hand, it does not confirm the potential benefits of the use of digital devices for poor comprehenders (Ruffini et al., 2023).

Consistent with our expectations, significant differences between task types (cloze vs. proof-reading) and lexical categories (functors, nouns and adjectives, verbs) emerged, indicating that task demands and lexical complexity interact in affecting recall. While in the cloze task, functors were generally easier to recall, possibly due to their higher frequency and syntactic predictability, contrasting results emerged in the proof-reading task. Nonetheless, these differences between lexical categories seem to vary according to age, with older students showing no differences in recalling verbs and functors. Moreover, age seems to play an important role in moderating the effect of comprehension abilities on comprehension scores, with reduced differences in older students between average and poor comprehenders. The interaction between the task type and lexical category suggests that lexical recall is context-dependent, and cognitive load and task structure could influence the retrieval of different lexical categories (Perfetti & Stafura, 2014).

4.2. Time Spent on Tasks Across Modalities and Groups (RQ2)

Although no significant differences emerged between groups of good, average, and poor comprehenders, students spent significantly more time on the proof-reading task than on the cloze task and in the digital modality compared to the paper one. These results could be an outcome of increased metacognitive demands during the comprehension-monitoring task, which requires not only reading and processing for comprehension but also proof-reading the text, evaluating each word and its plausibility in that context and editing words that are at odds with the expectations. Moreover, the cloze task was carried out on a summary of the previous text, therefore decreasing the time spent reading while completing the task. The absence of group differences in time spent on tasks suggests that all students, regardless of performance level, may engage similarly in time-consuming processes when faced with more complex tasks such as proof-reading, while this does not necessarily lead to equal accuracy performances. This could pinpoint a dissociation between effort and efficacy, especially for struggling readers (Clinton, 2019) and the results regarding the interaction among modality, task, and different levels of comprehension ability between groups could lead to further investigation to better understand the role of these variables in moderating the screen inferiority effect on time measures. Nonetheless, results showed that developmental differences might moderate the effect of the medium, and older students tend to overcome medium effects regarding the time spent on task.

4.3. Self-Evaluation and Calibration (RQ3)

The analysis of students' calibration regarding their own performance revealed no main effect of the performance group, indicating that metacognitive insight into task performance may not significantly vary across proficiency levels, in contrast with our expectations from the scientific literature about the difficulties of poor comprehenders in monitoring comprehension success. However, the significant effects of task and modality, as well as their interaction, suggest that calibration is task- and format-sensitive. Consistent with past research (Delgado et al., 2018; Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012), students exhibited greater calibration bias in the digital proof-reading task, indicating a higher degree of overconfidence. This finding supports the hypothesis

that digital environments may reduce metacognitive accuracy, particularly during complex tasks that require sustained monitoring and evaluation (Clinton, 2019). It also highlights the importance of considering the reading format not only in terms of performance outcomes, but in how it shapes students' perception of their own comprehension. Nonetheless, age seems to play an important role in moderating the effect of medium on calibration, and especially, older students with average comprehension abilities could overcome this bias over time. Further research is needed to address the reasons for these results that could highlight group-specific strategies when self-evaluating learning outcomes.

Altogether, these results help to highlight the circumstances of screen inferiority, with the emerging role of task demands and individual differences (comprehension abilities and age) in moderating its effects on comprehension process and performance. While the results of the study are of relevance to better understanding the individual differences at the base of the reading comprehension profiles across young students, there are several limitations to this research that should be addressed. First, the use of small narrative texts to sustain the comprehension tasks does not represent the typical learning content of these students. While the ecological aspect of the testing battery was a core aspect of our research in terms of task demands, new studies should involve different kinds of text (such as expository) while investigating lexical recalling across groups and modalities. Second, the use of the large lexical category of nouns and adjectives for the research, while it respects the lexical complexity of these kinds of stimuli, does not lead to disambiguating the role of these different components of the text during recalling and comprehension monitoring. Future studies could include a more specific differentiation between these categories. The differences in text length between the cloze-task and the proof-reading task, while it helped to express the differences in effort and efficacy between groups, should lead to careful consideration when addressing the results about time differences across tasks. Lastly, future studies could include longitudinal assessments of learning differences among good, average, and poor comprehenders when evaluated through digital and paper-based tasks targeting different dimensions of the comprehension process.

5. Conclusions

The present study provides new evidence suggesting that poor comprehenders do not differ from average and good comprehenders in digital versus paper reading tasks, although they do perform worse in both cloze and proofreading tasks. However, cloze and proof-reading tasks require more time in the digital modality, and the calibration bias is higher. Finally, nouns and adjectives are more difficult to retrieve compared to verbs and functors in proof-reading, whereas in the cloze task, functors are easier to be filled in compared to verbs, nouns, and adjectives.

These results have important implications, both practical and theoretical. First, they suggest that task design and lexical complexity should be carefully considered when assessing comprehension, particularly among students with varying levels of reading proficiency. Second, the influence of modality on metacognitive calibration highlights the need for instructional strategies that foster accurate self-monitoring, especially in digital contexts. Finally, the modality effect on time spent on the tasks suggests that time alone is not a reliable indicator of comprehension efficacy in a computer-based environment, emphasizing the need to think about other indicators of students' learning success that might be sensible to differences between groups of performances across modalities.

Future research should further explore the interplay among modality, task complexity, and metacognitive processes, possibly using eye-tracking or think-aloud protocols to unpack real-time cognitive effort. Moreover, longitudinal studies could examine how digital literacy and self-regulation skills evolve and influence comprehension over time.

Author Contributions: Conceptualization, P.B. and E.L.; methodology, P.B., E.L. and A.F.; validation, P.B., E.L. and A.F.; formal analysis, E.L.; investigation, A.F., S.G. and E.L.; resources, P.B.; data curation, A.F., E.L., S.G. and P.B.; writing—original draft preparation, E.L.; writing—review and editing, E.L. and P.B.; visualization, A.F. and S.G.; supervision, P.B.; funding acquisition, P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by ANASTASIS Soc. Cooperativa Sociale.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the University of Bologna Bioethical Committee approved the project (Prot. 0391784, 10 December 2024).

Informed Consent Statement: Informed consent was obtained from all parents involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy concerns.

Acknowledgments: We would like to thank the parents, teachers, and children who participated in data collection.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, 17(1), 18–32. [CrossRef]
- Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, 28(5), 1816–1828. [CrossRef]
- 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. [CrossRef]
- Ben-Yehudah, G., & Eshet-Alkalai, Y. (2020). Print versus digital reading comprehension tests: Does the congruency of study and test medium matter? *British Journal of Educational Technology*, 52(1), 426–440. [CrossRef]
- Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, 33, 73–79. [CrossRef]
- Bonifacci, P., Colombini, E., Marzocchi, M., Tobia, V., & Desideri, L. (2021). Text-to-speech applications to reduce mind wandering in students with dyslexia. *Journal of Computer Assisted Learning*, 38(2), 440–454. [CrossRef]
- Bonifacci, P., & Nori, R. (2016). *KBIT-2: Kaufman brief intelligence test: Standardizzazione italiana* (2nd ed.). Giunti O.S.
- Bonifacci, P., & Tobia, V. (2016). Crossing barriers: Profiles of reading and comprehension skills in early and late bilinguals, poor comprehenders, reading impaired, and typically developing children. *Learning and Individual Differences*, 47, 17–26. [CrossRef]
- Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, 76(4), 683–696. [CrossRef]
- Carretti, B., Borella, E., Cornoldi, C., & Bèni, R. D. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, 19(2), 246–251. [CrossRef]
- Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, 42(2), 288–325. [CrossRef]
- Cornoldi, C., Carretti, B., Cesaretto, J., & Viola, F. (2016). *Prove MT avanzate-3-clinica: La valutazione delle abilità di lettura e comprensione per la scuola primaria e secondaria di 1. grado: Manuale*. Giunti Edu.
- Cragg, L., & Nation, K. (2006). Exploring written narrative in children with poor reading comprehension. *Educational Psychology*, 26(1), 55–72. [CrossRef]
- 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. [CrossRef]
- 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. [CrossRef]
- Elwér, Å., Keenan, J., Olson, R., Byrne, B., & Samuelsson, S. (2013). Longitudinal stability and predictors of poor oral comprehenders and poor decoders. *Journal of Experimental Child Psychology*, 115(3), 497–516. [CrossRef]
- Florit, E., Carli, P., Lavelli, M., & Masón, L. (2022). Digital reading in beginner readers: Advantage or disadvantage for comprehension of narrative and informational linear texts? *Journal of Computer Assisted Learning*, 39(2), 432–445. [CrossRef]

- Furenes, M., Kucirkova, N., & Bus, A. (2021). A comparison of children's reading on paper versus screen: A meta-analysis. *Review of Educational Research*, 91(4), 483–517. [CrossRef]
- Georgiou, G., & Das, J. (2014). University students with poor reading comprehension. *Journal of Learning Disabilities*, 48(5), 535–545. [CrossRef] [PubMed]
- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116(2), 119–136. [CrossRef]
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10. [CrossRef]
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2(2), 127–160. [CrossRef]
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2nd ed.). AGS.
- Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris, & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71–92). Lawrence Erlbaum Associates Publishers.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394. [CrossRef]
- Lenhard, W., Schroeders, U., & Lenhard, A. (2017). Equivalence of screen versus print reading comprehension depends on task complexity and proficiency. *Discourse Processes*, 54, 427–445. [CrossRef]
- Mangen, A. (2016). The digitization of literary reading. *Orbis Litterarum*, 71(3), 240–262. [CrossRef]
- Mangen, A., Olivier, G., & Velay, J. (2019). Comparing comprehension of a long text read in print book and on kindle: Where in the text and when in the story? *Frontiers in Psychology*, 10, 38. [CrossRef]
- Mirandola, C., Ciriello, A., Gigli, M., & Cornoldi, C. (2018). Metacognitive monitoring of text comprehension: An investigation on postdictive judgments in typically developing children and children with reading comprehension difficulties. *Frontiers in Psychology*, 9, 2253. [CrossRef]
- Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, 39(1), 85–101. [CrossRef]
- Nation, K., & Snowling, M. J. (2008). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*, 69(4), 996–1011. [CrossRef]
- Peras, I., Mirazchiyski, E., Pavešić, B., & Recek, Ž. (2023). Digital versus paper reading: A systematic literature review on contemporary gaps according to gender, socioeconomic status, and rurality. *European Journal of Investigation in Health Psychology and Education*, 13(10), 1986–2005. [CrossRef]
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. [CrossRef]
- Ronconi, A., Veronesi, V., Mason, L., Manzione, L., Florit, E., Anmarkrud, Ø., & Bråten, I. (2022). Effects of reading medium on the processing, comprehension, and calibration of adolescent readers. *Computers & Education*, 185, 104520. [CrossRef]
- Ruffini, C., Tarchi, C., & Pecini, C. (2023). Which executive functions affect text comprehension and writing in paper and digital mode? An investigation in primary school children. *Computers & Education*, 207, 104936. [CrossRef]
- Salmerón, L., Vargas, C., Delgado, P., & Baron, N. (2022). Relation between digital tool practices in the language arts classroom and reading comprehension scores. *Reading and Writing*, 36(1), 175–194. [CrossRef] [PubMed]
- Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, 51, 61–73. [CrossRef]
- Singer, L., & Alexander, P. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research*, 87(6), 1007–1041. [CrossRef]
- The Jamovi Project. (2024). *jamovi* (Version 2.6) [Computer Software]. Available online: <https://www.jamovi.org> (accessed on 26 June 2025).
- Tobia, V., & Bonifacci, P. (2020). Look back at text or rely on memory? Efficacy of reading comprehension strategies in good and poor oral comprehenders. *Journal of Research in Reading*, 43(4), 536–555. [CrossRef]
- Tong, X., Deacon, S., & Cain, K. (2013). Morphological and syntactic awareness in poor comprehenders. *Journal of Learning Disabilities*, 47(1), 22–33. [CrossRef]
- Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge University Press.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.