

Article

The Text-Belief Consistency Effect Among Recent Upper Secondary Graduates: An Eye Tracking Study

Mariola Giménez-Salvador ^{*}, Ignacio Máñez  and Raquel Cerdán 

ERI Lectura and Department of Developmental and Educational Psychology, Faculty of Psychology and Speech Therapy, University of Valencia, 46010 Valencia, Spain; ignacio.manez@uv.es (I.M.); raquel.cerdan@uv.es (R.C.)

^{*} Correspondence: magisal2@alumni.uv.es

Abstract: Readers tend to allocate more cognitive resources to processing belief-consistent than belief-inconsistent information when reading multiple texts displaying discrepant views. This phenomenon, known as the text-belief consistency effect, results in individuals being more prone to making biased decisions and falling victim to manipulation and misinformation. This issue is gaining relevance due to the undeniably vast amount of information surrounding us. Hence, schools must ensure that students complete their education prepared to face this challenge. However, international surveys and research indicate a generalized literacy deficiency among students. In the present study, recent upper secondary graduates read four texts discussing a controversial topic to explore whether they effectively overcome the text-belief consistency effect or not. Eye tracking was used to explore immediate (or passive) and delayed (or strategic) text processing, and an essay task was used to measure their resulting mental representation of the text content. Results revealed no significant differences in immediate and delayed processing depending on whether the arguments were belief-consistent or belief-inconsistent. Moreover, essays displayed a balanced and unbiased approach to the debate. Despite these results suggesting this population may be capable of overcoming the text-belief consistency effect, limitations in the study and alternative explanations must be explored before drawing definite conclusions.

Keywords: reading comprehension; cognitive processes; controversies; beliefs; argumentation; eye tracking; high school; university



Citation: Giménez-Salvador, M.; Máñez, I.; Cerdán, R. The Text-Belief Consistency Effect Among Recent Upper Secondary Graduates: An Eye Tracking Study. *Educ. Sci.* **2024**, *14*, 1259. <https://doi.org/10.3390/educsci14111259>

Academic Editor: Myint Swe Khine

Received: 7 October 2024

Revised: 29 October 2024

Accepted: 6 November 2024

Published: 18 November 2024



Copyright: © 2024 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 development and growth of the World Wide Web has created an easy and fast way to access a vast amount of information beyond measure on any potential topic of interest. To satisfy individual needs and to learn and make informed decisions on everyday life matters (e.g., global warming, vaccination, ongoing wars), people need to understand information across diverse sources. Furthermore, students at schools or universities are required to complete tasks involving the navigation of several texts and the integration of relevant ideas. Consequently, reading multiple documents, sometimes contradictory, has become an everyday activity for both informal and formal learning contexts [1].

Making an effective use of multiple sources of information is a challenge that requires advanced reading literacy skills that go beyond mere word decoding and passage comprehension [2]. Since being able to master these skills is essential for higher education and adult life, educational systems must ensure that students acquire them before leaving school. However, international sample surveys reflect that there is significant potential for improvement in many countries, including Spain. For instance, the Programme for International Student Assessment (PISA), promoted by the Organisation for Economic Co-operation and Development (OECD), evaluates the proficiency of students in reading, mathematics and science as they near the end of their basic compulsory education at around 15 years old. The results of the survey conducted in 2022 indicate that reading proficiency in

Spain does not significantly differ from the OECD and European Union means [3]. Despite this, the score obtained is among the lowest since the first PISA report in 2000. Another OECD assessment survey is the Programme for the International Assessment of Adult Competencies (PIAAC). This survey evaluates literacy, numeracy, and problem-solving skills among adults of different socio-demographic characteristics. In this case, according to the outcomes of the last assessment in which Spain participated, conducted from 2011 to 2012, recent upper secondary graduates (16–19 years old) in Spain scored a bit below the OECD mean [4].

Therefore, the results from both surveys highlight the necessity of investing efforts to deepen the understanding of the complex processes involved in reading and address the most common challenges faced by individuals. One of these difficulties arises when encountering multiple documents that present inconsistent or conflicting information, as the reader must be aware of the conflict and manage it [5]. In these circumstances, there is the potential risk that the readers' prior beliefs take a main role in guiding the processing of the information, in such a way that the processing results in a one-sided and biased mental representation of the dispute [6]. Specifically, belief-consistent information holds an advantage in processing compared to belief-inconsistent information. This phenomenon is called the text-belief consistency effect (e.g., [7–9]) and has been specifically explored among university students (for a systematic review, ref. [10]) and adolescents (e.g., [11]). Overcoming this effect results in the construction of an integrated and balanced mental model, that is, a mental representation reflecting both the overlapping and divergent information found across the texts, as well as the alternative positions presented in the debate [12].

However, research on the text-belief consistency effect, as well as PISA and PIAAC assessment surveys, primarily utilizes performance data to measure reading literacy, which constitute a non-immediate (offline) type of information. In recent years, eye tracking has emerged as a significant tool in the field of educational science [13]. *Eye tracking* is a technique that offers real-time (online), nonintrusive, and objective data on information processing during reading. This methodology can supply details at a microscopic level, providing information about how individuals allocate their visual attention and the strategies they employ while reading, even during unconscious processes [14]. Therefore, eye tracking can provide unique insights that can be triangulated with other types of information, such as performance data or verbal data, e.g., through think-aloud protocols, or interviews [15]. Despite its numerous advantages, there is limited research using eye tracking to study comprehension of multiple documents, in specific, the text-belief consistency effect (e.g., [16,17]).

The present study addresses this gap of knowledge. Through an empirical study using eye tracking and an essay task, it is expected to shed light on how students who have just completed upper secondary education manage information from contradictory sources. Results will contribute to constructing a more detailed diagnosis of the skills of this specific population. In turn, the eventual clarification of this matter will encourage teachers, educational guidance counselors, researchers, policymakers, and other educational stakeholders to identify effective practices for shaping the appropriate cognitive processes when reading contradictory multiple texts. Furthermore, the data collected in the present study will elucidate the utility of eye tracking methodology for researching the comprehension of multiple conflicting documents.

1.1. Comprehension of Multiple Texts: A Complex Endeavor

The conventional conception of reading encompasses two main concepts: firstly, decoding written words, sentences, or texts, and secondly, comprehending the meaning of the written text. This aligns with what is known as the *simple view of reading* [18], primarily due to the exclusion of various other skills and factors that also contribute to the reading experience. Despite its limitations, the simple view of reading acknowledges the fundamental role of its components in every reading process and has served as a foundation for

the development of new research and enriched models and theories, especially encouraged by the introduction of digital technologies in reading [2,19,20].

A comprehensive compilation of the latest empirical and theoretical contributions is represented in the recently revised framework for assessing reading literacy elaborated by PISA 2018 [21]. PISA 2018 states that, to function in our actual knowledge-based society, it is not enough to memorize and accumulate information from independent texts. Instead, it becomes essential to master reading skills such as finding, selecting, interpreting, integrating, and evaluating information from multiple texts. Individuals are expected to engage in these processes in order to achieve their goals, which will vary throughout their lives, including academic or work-related purposes, personal aspirations, and desires to participate in the political, economic, or cultural life of today's society. Based on this statement, PISA 2018 proposes a reading framework of processes consisting of two main components and their corresponding subcomponents [21]: text processing (which includes reading fluently, locating information, understanding, and evaluating and reflecting) and task management (involving setting goals and plans and monitoring or regulating). Some components of this framework will be now reviewed, particularly the higher-order processes, and additional theoretical contributions will be incorporated when appropriate.

With respect to the process of understanding (also called comprehending), PISA 2018 presents Kintsch's idea of a situation model [22]. A *situation model* is a mental representation of the meaning of the text resulting from the generation of inferences. These inferences involve the connection of different parts of the text (bridging inferences) or the connection of the text with the reader's prior knowledge (elaborative inferences), in order to create spatial, temporal, causal, or other kinds of associations [22–24]. However, Kintsch's situation model does not consider how readers construct meaning from more than one text. In this context, the documents model is proposed [6,25,26]. The *documents model* requires the construction of two structures, an integrated model and an intertext model. When readers encounter multiple texts, a situation model is created for each of them. These models are then linked together to construct an *integrated model*. The more similar the content of the texts is, the more merged the situation models become within the integrated model.

The integrated model is then complemented by the *intertext model*, which represents the source of each text (e.g., author, context, type of document, communicative intentions), the associations between the sources and the content (e.g., author A asserts X), and the connections between the different sources (e.g., author A agrees or disagrees with author B). This last component is key to understanding how coherence can be achieved in a reader's mental representation despite documents presenting inconsistent perspectives. Although contradictions across texts prevent individuals from merging documents' content into a single unified mental model, incorporating source information into the mental representation offers a solution to this coherence issue. However, building a documents model may not be enough, as the conflict between the sources is not being solved. That is why, to achieve further coherence, readers should evaluate the validity of the claims and the credibility of the sources [5,10].

Finally, in relation to task management, a framework that considers the role of goals in text processing is the reading as problem solving (RESOLV) model [2,19]. The RESOLV model proposes two distinct constructs involved in reading: a context model and a task model [2]. Firstly, readers construct a *context model*, that is, a mental representation including the individual's interpretation of any element from the physical and social context that they can access and consider relevant. The potential contextual features regarded by the reader include characteristics of the request, the requester, the audience, support and obstacles, and the self. With respect to the *self* dimension, it includes the reader's perception about themselves as readers, e.g., their prior knowledge or beliefs about the matter approached by the texts. Based on the context model, the reader constructs the *task model*, a mental representation of the goal to be fulfilled, that is, the subjective interpretation of the outcome that should be achieved and the means available to reach it. In turn, this task model drives

readers' choices and strategies concerning what and how to read. For example, depending on the goal, individuals may decide to read a specific source, text, or passage and skim or skip the rest. Additionally, readers may opt for deeper information processing, including generating bridging or elaborative inferences, integrating ideas coming from different texts, representing the source information, or evaluating the validity of the claims—or, conversely, they may engage in more superficial processes. The task model is updated as the individuals progress in their reading and assess whether they have already achieved a satisfactory reading outcome or not. This assessment and updating reflect the readers' monitoring (or regulating, metacognitive) processes.

Considering that, according to the RESOLV model [2,19], prior beliefs can influence the reader's task model and the evaluation of the reading outcome, it should not be unexpected that, when reading controversial matters, the prior position on the debate influences the reader's management of the conflict to some extent. The Sections 1.2 and 1.3 will explore how prior beliefs affect comprehension of contradictory documents and will discuss the measures that have allowed the observation of this phenomenon.

1.2. The Role of Readers' Prior Beliefs: The Text-Belief Consistency Effect

When reading conflicting texts, there is the risk that the reader's prior beliefs influence information processing. Specifically, readers may tend to favor belief-consistent information by allocating more cognitive resources to its processing while disregarding belief-inconsistent information. This phenomenon has been named the text-belief consistency effect (e.g., [7,9]) and can be considered a variant of the confirmation bias [27] or selective exposure effect [28]. To contextualize this effect within a more comprehensive model of how readers evaluate the validity of information and its effects on comprehension, Richter and Maier propose the two-step model of validation [10].

In this model, the authors suggest that readers engage in one or two steps of validation when processing conflicting information. While reading, individuals may encounter textual elements (e.g., words, sentences) that implicitly retrieve and activate information from their long-term memory into their working memory, such as their prior beliefs. Once these beliefs have been triggered, they serve as a heuristic to validate textual input, which means that individuals conduct an involuntary check of the plausibility of the text information based on these beliefs. The information that is not considered plausible is that which is inconsistent with their prior beliefs, implying that this information is less likely to be integrated into the situation model readers are constructing as they read. Consequently, readers create a one-sided representation of the debate. This process of validation corresponds to the first step, *routine validation*, which occurs in an unconscious and passive manner, while the bias towards belief-consistent information is precisely the text-belief consistency effect.

After engaging in routine validation, there is the possibility that readers perceive a lack of coherence in their mental representation due to the detection of great inconsistencies between their beliefs and the text content. If readers are not satisfied with this reading outcome, they may proceed to *strategic validation* (second step), consisting in the elaboration of the belief-inconsistent information in order to resolve the conflict. Therefore, this second step reflects that if readers can engage in metacognitive control mechanisms, they may become aware of their biases, prompting them to seek out additional information. Furthermore, these two steps align well with the dual-process theory of reasoning (e.g., [29]), which posits that individuals often rely on intuitive, automatic responses before engaging in more deliberate, analytical thought.

Moreover, regarding strategic validation, the authors suggest that, as this deeper processing of the belief-inconsistent information is a demanding task, it occurs only when readers' goals require it, which may be related to how cognitively capable and motivated they feel or to the instructions given by an external source, among other factors. Therefore, if the conditions are unfavorable, readers will remain engaged in routine validation. However, strategic validation offers great benefits for comprehension and memory, resulting in a balanced and coherent mental model. This notion can be related to cognitive load

theory [30]. According to this model, when individuals need to complete a task, there are two kinds of cognitive load to consider: intrinsic cognitive load, associated with the processing and learning of information, and extraneous cognitive load, associated with the instructional methods employed. To manage both types of cognitive load, working memory resources are required, which are limited. Therefore, when conditions reduce the working memory capacity (e.g., low motivation, tiredness) or when extraneous cognitive load is high (e.g., ambiguous instructions), fewer resources remain available for processing and learning the texts. In these cases, individuals may prefer to maintain simpler strategies (i.e., routine validation) and avoid engaging in deeper processing (i.e., strategic validation).

Ample research findings provide evidence for the existence of the text-belief consistency effect and the two steps of validation (for a systematic review, ref. [10]). For instance, Stadtler et al. studied the impact of different task instructions on the quality of the essays written by university students after reading texts about a controversial topic [1]. Specifically, the authors compared asking students to write an argument, a summary, or a list of keywords. The results revealed that only the participants in the argument condition included a two-sided approach to the dispute in the essay. This could mean that this type of instruction contributes to the creation of a reading goal promoting strategic validation of the belief-inconsistent information, while the other two instructions keep readers engaged in routine validation. Other studies have directly assessed prior beliefs before the reading episode. In these investigations, it has been found that both secondary school and university students, in normal circumstances, elaborate biased essays favoring belief-consistent information [31–34], demonstrate enhanced memory for belief-consistent information in a recognition task [7,9,11,35], and spend more time reading belief-consistent texts, as indicated by a think-aloud protocol [36].

These results can be further supported with evidence gathered through research that employs eye tracking methodology. For example, it has been found that when readers first encounter inconsistent information their reading times increase, even when they are not given the instruction to assess the consistency of the information (e.g., [37–39]). This finding may signal that a process of monitoring is occurring during initial reading, which would correspond to routine validation. However, this boost in reading times does not indicate that readers are actively trying to understand inconsistent information, but rather suggests a mere disruption in comprehension. Thus, Richter and Maier propose that, despite this disruption, readers will continue through the text without attempting to resolve the conflict, as predicted for routine validation [10]. Furthermore, other eye tracking studies have found an association between revisiting previously read belief-inconsistent sentences and the action of seeking information to resolve the conflict (e.g., [40]), a process that corresponds to strategic validation.

Despite eye tracking being a great methodology for accessing readers' online cognitive processes, there are not many studies conducted with the aim of confirming the existence of the text-belief consistency effect. In the Section 1.3, an overview of eye tracking methodology will be presented, with an emphasis on the links found between cognitive processes and gaze measures, particularly when reading multiple contradictory documents.

1.3. Associations Between Gaze Data and Cognitive Processes When Reading Multiple Texts

Over the past four decades, researchers in the field of educational science have extensively employed eye tracking measures to understand the processes underlying reading comprehension [41]. That is because eye tracking technology offers several benefits over other types of instruments. Unlike offline techniques (e.g., questionnaires, essays, or interviews), eye trackers can gather direct and objective data reflecting even the most unconscious and automatic cognitive processes during reading [42]. Certainly, the functionality of eye tracking technology lies in the assumption that there is a correspondence between the eye movements captured by the device and the underlying cognitive processes. This assumption is commonly known as the *eye mind hypothesis* [43,44]. Specifically, eye movements can be associated with changes of visual attention [43], depth of information

processing [14], or difficulty in processing [45]. Eye tracking technology provides two main types of measurements: fixations and saccades. *Fixations* correspond to the brief pauses of the eye on a specific location, while *saccades* are the rapid movements of the eye from one fixation point to another. Both fixations and saccades can be reflected through different specific measurements. Lai et al. classify these measurements into three categories: temporal, spatial, and count [46]. Firstly, the *temporal* category refers to the duration of an eye movement measurement (e.g., total fixation duration, that is, the sum of the duration of all fixations on a specific area). Secondly, the *spatial* category alludes to the locations, directions, distances, or sequences of fixations and saccades (e.g., fixation sequence). Finally, the *count* category encompasses measurements indicating the frequency of eye movements (e.g., total fixation count, that is, the number of fixations on a specific area).

The majority of eye tracking studies in reading comprehension have focused on isolated words or sentences; however, these do not represent real-world reading tasks [47]. Over the past 20 years, there have been various efforts to broaden the use of eye tracking to investigate reading within more complex contexts, such as reading multiple documents of a certain length, particularly texts presenting discrepant perspectives [41]. In these studies, cognitive scientists typically differentiate between early (or initial) and late (or delayed) text-processing measures [14,48–50]. Early measures include measurements such as *first fixation duration*, which refers to the duration of the first fixation on a specific area (with “area” interpreted as a sentence in this context), *first-pass reading duration*, which is the total duration of all fixations on a sentence during the first pass (i.e., the first encounter of the reader with the sentence before moving on to the next one), and *first-pass re-reading duration*, which represents the total duration of those fixations going back to reinspect parts of a sentence that the reader has not yet finished reading entirely. In sum, these measures signal the readers’ behavior during the first time they encounter a specific area (i.e., a sentence). However, readers’ gaze sometimes regresses or returns to a part of the text they have previously read. This process is represented by measures such as *lookback reading duration*, which is the total duration of all fixations on a sentence after the first pass, and *lookback count*, which refers to the number of fixations on a sentence after the first pass. These measures indicate delayed stages of text processing and suggest an interruption to the typical reading flow, such as encountering an unfamiliar word, having difficulty in grasping the meaning of a passage, or even struggling to create a coherent mental representation of the text content.

Most of the studies employing multiple contradictory documents have focused on analyzing eye movements in relation to source information (e.g., [51–53]), the whole text (e.g., [52]), different sections of the text (e.g., [54]), or comparing conflicting and nonconflicting passages (e.g., [55]). However, there are only a few studies that have employed eye tracking methodology to investigate the text-belief consistency effect. To the best of our knowledge, the only two studies that have done so are those by Maier et al. [16] and Abendroth and Richter [17], both of which included university students in their sample of participants. Maier et al. [16] identified an increase in participants’ first-pass re-reading times for belief-inconsistent information, aligning with the common increase in reading times found during routine validation signaling a disruption in comprehension (e.g., [37–39]). Furthermore, these slow-downs were more pronounced for those participants with stronger prior beliefs, supporting Voss et al.’s findings [56]. The authors also found a higher lookback count for belief-inconsistent information when presenting the texts in an alternating manner (e.g., first reading a belief-consistent document, followed by a belief-inconsistent document before reading again a belief-consistent text), in contrast to when texts were presented in a blocked manner (i.e., individuals read documents of the same belief type successively). Based on Rinck et al.’s [40] conclusion that revisits to belief-inconsistent information are associated with attempts to solve the conflict, these findings may suggest that an alternating mode of presentation contributes to a heightened awareness of inconsistencies in the reader’s mental representation. This heightened awareness, in turn, may promote the creation of a reading goal that encourages strategic

validation of belief-inconsistent sentences [7,9]. The authors also investigated the text-belief consistency effect using performance data, specifically, an essay. Aligning with previous findings, participants generally wrote more belief-consistent arguments in their products, although this effect was reduced if texts were presented alternately.

Regarding the second mentioned study, Abendroth and Richter [17] found very similar results. Firstly, they also observed an increase in reading times when reading belief-inconsistent information. However, the eye tracking measurement used in this case was first-pass reading times instead of first-pass re-reading times. Secondly, the authors found a longer lookback reading duration for belief-consistent information when the instruction suggested participants take a belief-consistent viewpoint, whereas similar lookback times were noted for both types of information when the instruction suggested that they take a belief-inconsistent viewpoint. These observations could indicate that the belief-consistent viewpoint, as it is in line with the readers' predisposition to favor this type of information, results in a deeper processing of belief-consistent ideas. In contrast, when readers are assigned a belief-inconsistent viewpoint, instead of only favoring this type of information, they deeply process both. Therefore, these outcomes also align with the two-step model of validation [10]. Furthermore, as in the prior study, a performance task was also introduced, in this case, a recognition task. Coherently with the aforementioned results, there was enhanced memory for belief-consistent information within the belief-consistent reading perspective, while memory remained balanced within the belief-inconsistent reading perspective.

1.4. The Current Study: Objectives, Research Questions, and Hypotheses

The goal of the present research is to explore the extent to which recent upper secondary graduates' prior beliefs on a controversial topic are related to their reading behavior (eye gaze data) and their positioning in a written essay. In other words, the so-called text-belief consistency effect is examined. Although this specific population should be prepared for higher education and adult life, data from international surveys indicate a deficiency in their literacy competence [3,4]. Nowadays, mastery of multiple text integration, particularly if the texts are contradictory, is essential due to the vast amount of available information [57]. Thus, as commented, acquiring more details about these students' patterns of reading would help in making a complete diagnosis of their abilities and would encourage educational stakeholders to devise solutions and implement them. As previously mentioned, prior studies have observed that beliefs affect both the processing and performance of secondary school and university students when reading multiple contradictory documents. In particular, belief-consistent information has been shown to have an advantage over belief-inconsistent information (see Richter and Maier's review [10]). Different instruments have revealed this pattern, including performance-based tools such as essays [16,31–34] and recognition tasks [7,9,11,17,35], as well as process-based tools like think-aloud protocols [36] and eye tracking technology [16,17]. Employing both types of instruments at the same time can offer a comprehensive understanding of readers' management of this challenging task. That is why the following research questions are formulated for the present study. Research question 1: How do prior beliefs affect eye movements while reading multiple conflicting texts? Research question 2: How do prior beliefs affect argument integration (essay performance) after reading multiple conflicting texts?

As evidenced by the number of studies just referenced, while research employing essays is extensive, eye tracking remains underdeveloped. Eye tracking technology offers significant strengths, as it provides real-time and objective information about the microscopic and unconscious aspects of reading [14]. Thus, it can greatly benefit this area of research, especially when complemented with other tools, as in the present study. Therefore, based on previous investigations, the following hypotheses are proposed. Regarding the first research question, readers' eye movements, including first-pass reading duration and lookback reading duration, will vary depending on the stance of the arguments (in favor of or against the controversial topic). Specifically, a disruption in comprehension, typically seen during routine validation will be indicated by a longer first-pass reading duration for

arguments that conflict with prior beliefs (Hypothesis 1). This would suggest that prior beliefs have been activated and that the discrepancies between them and the text arguments have been detected. This hypothesis is directly supported by Abendroth and Richter's study [17]; however, Maier et al. [16] did not find an increase in first-pass reading duration, but rather in first-pass re-reading duration. The authors suggest that the former index is not sensitive to inconsistency effects, aligning with Rinck et al.'s findings [40].

Furthermore, if the text-belief consistency effect is not overcome, a greater allocation of cognitive resources to arguments coherent with prior beliefs will be reflected by a longer lookback reading duration for these arguments, indicating a lack of engagement in strategic validation for belief-inconsistent arguments (Hypothesis 2). This hypothesis is supported by the studies of both Abendroth and Richter [17] and Maier et al. [16]. In relation to the second research question, it is expected that a relationship will be observed between prior beliefs and the essay's positioning (i.e., arguments in favor or arguments against). Similar to Hypothesis 2, if the text-belief consistency effect is not overcome, participants' positioning in the written essay will be coherent with their prior beliefs, signaling that participants have created an unbalanced mental model of the texts favoring belief-consistent arguments (Hypothesis 3). For this hypothesis, there is substantial evidence supporting it [16,31–34], especially in comparison to the previous two.

2. Materials and Methods

2.1. Participants and Design

Twenty-two undergraduate students enrolled in their first year of the bachelor's degree in Primary Education Teacher Training at a Spanish university volunteered to participate in the study in exchange for class credit. However, three students could not participate, either due to schedule incompatibilities ($n = 2$) or to visual acuity problems ($n = 1$). Furthermore, two students were excluded from the analysis due to poor registration of their eye movements. The remaining seventeen participants were on average 18.41 years old ($SD = 0.60$). All of them were women, native Spanish speakers, had a high level of English proficiency, and their degree access score was moderately high on average ($M = 11.68$, $SD = 0.69$). Regarding the study's methodology, it is quantitative. Specifically, both research questions employed an observational (correlational) design, as the focus was on examining the relationships between various continuous variables (prior beliefs and gaze measures for the first research question and prior beliefs and essay's positioning for the second research question). However, the first research question also incorporated a repeated measures within-subjects design to account for the argument stance (in favor or against) when examining the relationship between prior beliefs and the gaze measures. The order of presentation of the texts (and consequently, arguments) was controlled by displaying all texts randomly to the participants. Therefore, the study combined both observational and experimental elements.

2.2. Materials

Each participant read four conflicting texts on the homework debate written in English. Two of the four texts discussed the pros of assigning homework as a teaching method, whereas the other two discussed the cons. Therefore, each of the four texts included sentences considered as arguments (in favor or against, depending on the text stance), while the other sentences were considered as irrelevant. The texts were extracted from the materials used in a previous paper-based study [58]. In that study, the texts had been adapted from online newspapers, and the arguments had been identified with an inter-rater agreement of over 90%. However, to collect high-quality gaze data in the present study, the four documents were slightly modified to comply with the characteristics of the eye tracking device and software, which are explained in the Section 2.3. The criteria used to modify the original texts were as follows: (a) attaining a word count for each text ranging between 200 and 250 words, (b) maintaining the same arguments as in the original study, (c) keeping arguments separated from each other and confined to a point-to-point sentence,

(d) reducing overlapping between arguments, and (d) ensuring an equal total number of arguments in favor and against.

Ultimately, the documents were limited to 220–226 words (see Table 1), roughly the same number of words. The texts included a total of 10 sentences considered as arguments in favor and 10 sentences considered as arguments against (see Tables 2 and 3). The complete documents can be found in Appendix A, as well as the arguments, which are highlighted in grey. The readability scores of the four texts were computed using the Flesch reading-ease score (FRES) test to denote how difficult a text written in English is to comprehend [59]. The readability formula is the following: $206.835 - 1.015 (\text{total words} \div \text{total sentences}) - 84.6 (\text{total syllables} \div \text{total words})$. The FRES scale distinguishes eight levels of text difficulty, ranging from extremely difficult to read (<10.00; comprehended by university graduates) to *very easy to read* (90.00–100.00; understood by primary school students). The mean readability score of these four documents was 48.88 ($SD = 3.46$), which corresponds to the category *difficult to read*, indicating that the texts were adequate for undergraduate students, as they were readable enough, yet required some effort to be comprehended. The scores for each document were fairly similar, as Table 1 shows.

Table 1. Overview of the reading material.

Title of the Text	Word Count	Readability Score	Text Stance
"Too Much Homework? Some Parents Are Just Opting Out"	220	54.25	Against
"Homework: Supporting Student Success and Growth"	226	44.70	In favor
"A Primary School Teacher Stopped Assigning Homework"	223	47.63	Against
"The Homework Debate: How Homework Benefits Students"	223	48.93	In favor

Table 2. List of arguments against.

Title of the Text	Arguments Against
"Too Much Homework? Some Parents Are Just Opting Out"	• Parents' conflict between homework and free time. • Lack of free time caused by homework. • Popularity of the idea of banning homework. • Recommendation: balance with an age-appropriate level of academic engagement. • Recommendation: talking to teachers to adjust demands.
"A Primary School Teacher Stopped Assigning Homework"	• Elementary school teacher's opinion: no more homework. • No proven effect on performance. • Success correlates with free time. • Variables to consider in the debate (age, type of homework, time). • Recommendation: no heavy homework loads.

Table 3. List of arguments in favor.

Title of the Text	Arguments in Favor
"Homework: Supporting Student Success and Growth"	• Homework as an opportunity to interact. • It promotes perseverance. • It promotes self-esteem. • Link between homework and performance in secondary school. • Useful for primary-school students to prepare them for secondary school. • Recommendation: homework with a clear goal and purpose.
"The Homework Debate: How Homework Benefits Students"	• Homework promotes organization. • It promotes responsibility. • It gives students a chance to review. • It allows teacher supervision and feedback.

2.3. Eye Tracking Equipment

Tobii Pro Nano was used to track eye gaze (i.e., eye movements). This device is a small, portable, and non-invasive eye tracker that registers data binocularly at a sampling rate of 60 Hz. It was mounted onto a 17-inch (43.18 cm) IPS laptop with a maximum resolution of 2560×1600 pixels (16:10 format). Tobii Pro Lab software (lab version 1.207) was used to design the reading material, record the data, and obtain the indices of visual behavior. An important aspect to consider is the delimitation of the *areas of interest* (AOIs), which are regions in stimuli that enable numerical and statistical analysis [60]. In this case, the AOIs from which information was to be extracted were the arguments in favor of or against homework found within the four texts to be read. While AOI can be manually created, Tobii Pro Lab offers the option to automatically generate AOI for characters, words, or sentences. In this study, the AOI were created at the sentence level.

2.4. Measures

2.4.1. Prior Beliefs

Prior beliefs on the homework debate were measured using six items in a 5-point Likert-type scale (with 1 indicating *Strongly disagree* and 5 *Strongly agree*) already used in a previous study [58]. The questions were the following: (a) “I am quite familiar with the current debate about homework”, (b) “I am quite interested in reading texts dealing with a debate about homework”, (c) “I believe homework is completely necessary for school learning”, (d) “I believe homework should be eliminated from schools (children should work during school time)”, (e) “I think homework can help children learn better”, and (f) “I think homework does not help the teacher in supporting the students’ learning processes”. After reversing scores from the negative items (d, f), the reliability coefficient measured using Cronbach’s alpha was found to be 0.606. Removal of item b substantially improved the coefficient to 0.800. Scores were added and averaged to create a 5-point single-belief score, ranging from 1 (*completely against*) to 5 (*completely in favor*).

2.4.2. Gaze Measures

Two indices of visual behavior were obtained for each AOI (argument): (a) the total duration of fixations during the first pass inside the AOI, meaning the sum of the duration of all fixations on the previously unread argument before moving on to the following sentences (*first-pass reading duration*), and (b) the total duration of fixations after the first pass inside the AOI, that is, the sum of the duration of all regressive fixations to this previously read argument (*lookback reading duration*). To account for variations in sentence length for each AOI, the two indices were adjusted based on the number of words in each sentence [16,17]. After that, the arithmetic mean was calculated for each of the adjusted measures across the two types of AOIs. Therefore, each participant had one aggregated measure per time parameter for arguments in favor and arguments against, resulting in a total of four eye movement dependent variables (in msec/word): (a) first-pass reading duration for arguments in favor, (b) first-pass reading duration for arguments against, (c) lookback reading duration for arguments in favor, and (d) lookback reading duration for arguments against.

As previously mentioned, the rationale for selecting these gaze metrics is based on previous studies that identified a relationship between the processes occurring at each step of the two-step model of validation [10] and specific eye movement behaviors. First, as mentioned previously, studies have observed an increase in reading times when readers first encounter inconsistent information [37–39], a process associated with the first step, routine validation. Particularly, Maier et al. [16] found an increase in first-pass re-reading times, while Abendroth and Richter [17] observed an increase in first-pass reading times. Since Tobii Pro Lab only provides first-pass reading times and not first-pass re-reading times, this metric was selected to represent routine validation. Second, research has found an association between revisiting previously read belief-inconsistent sentences and the process of seeking information to resolve the conflict [40], a behavior corresponding to strategic

validation, the second step. Specifically, Maier et al. [16] identified this relationship using the metric of lookback count, while Abendroth and Richter [17] used lookback reading duration. As Tobii Pro Lab provides only lookback reading duration and not lookback count, this metric was selected to represent strategic validation.

To ensure data quality, a nine-point binocular calibration and a four-point validation of the calibration were used prior to the start of the reading episode. Accuracy and precision indices, as well as a data loss percentage, were obtained for both eyes after this calibration-validation procedure. The lower the values for accuracy, precision, and data loss are, the better. As Figure 1 shows, while *accuracy* refers to the average disparity between the actual gaze position and the recorded gaze position, *precision* reflects the ability of the eye tracker to reliably replicate identical gaze point measurements from one sample to another [13]. The percentage of *data loss*, on the other hand, represents the amount of eye tracking samples that are not correctly identified [60]. It is typical to lose samples due to blinking or looking away from the monitor, but other reasons can include the lighting conditions, wearing glasses, or technological challenges [13]. The online platform Tobii Connect, for example, uses an accuracy threshold of $\leq 0.5^\circ$ deviation and a precision threshold of $\leq 0.2^\circ$ deviation [61]. Regarding data loss, the Tobii Pro Lab User Manual [60] indicates that approximately 10% of data loss can be expected, primarily due to blinking. Thus, these were the guidelines followed in the study.

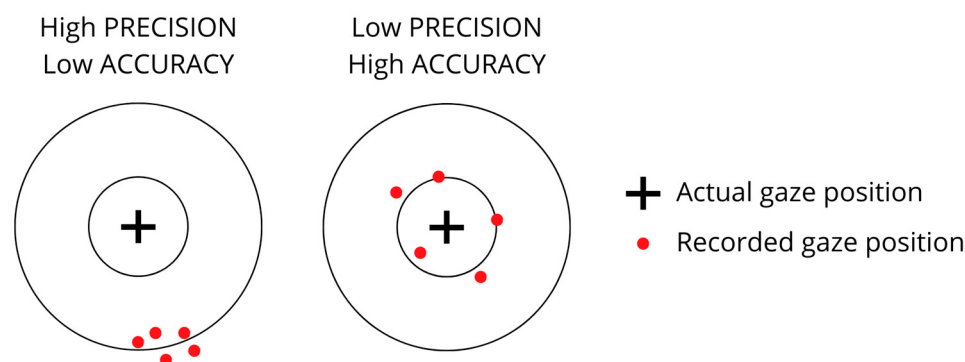


Figure 1. Visual representation of accuracy and precision.

2.4.3. Argumentative Essay Scores

Participants were instructed to write a 300 to 500-word essay in English on the homework debate using the same laptop they used for reading the texts. The following instruction was provided: “Do you think assigning homework is a useful teaching method? Please support your conclusions based on the arguments expressed in the texts”. Participants did not have access to the four texts while they were writing the essay. Furthermore, responses were analyzed based on the identification of participants’ inclusion of text arguments (refer to Tables 2 and 3). As commented, the four texts used in the present study were extracted and adapted from a previous paper-based study [58], in which arguments had been identified with an inter-rater agreement of over 90%. Consequently, three measures were obtained: sum of arguments in favor (ranging from 0 to 10), sum of arguments against (ranging from 0 to 10), and a positioning score, i.e., the difference between the two aforementioned indices, ranging from -10 (*completely against*) to 10 (*completely in favor*).

2.5. Procedure

Data collection took place in two phases. First, participants filled a 5 min questionnaire through Google Forms in which they gave their informed consent, were asked demographic questions, reported whether they had learning difficulties, vision problems or used glasses incompatible with the eye tracker, and answered the six items about their prior beliefs on the homework debate (see Appendix B). After excluding a participant who did not meet the inclusion criteria, the eligible subjects were sent a calendar by email displaying the

available individual sessions from which they could choose. Two participants left the study due to schedule incompatibilities.

Second, the remaining participants completed the reading task and wrote the essay in an individual session in the laboratory. Participants were informed about the meeting point and were advised not to wear eye makeup and to tie their hair back in a manner that did not cover their eyes. If they normally wore glasses or contact lenses for reading, they were asked to bring them. In the case of wearing glasses, they were provided with wipes to clean them before performing the task. The individual sessions lasted a maximum of 60 min. However, these sessions required a prior preparation of the eye tracking equipment and environmental conditions. On the one hand, Tobii Pro Lab needed to be downloaded and installed, and Tobii Pro Nano had to be mounted on the laptop, configured with the screen and tested to troubleshoot any possible errors. On the other hand, the workspace had to be arranged, including the seating (chair and position), work surface, lighting conditions, and potential distractors (Figure 2).

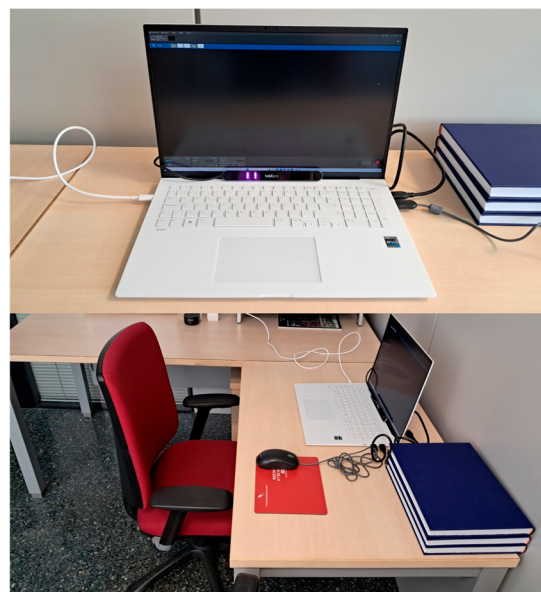


Figure 2. Workspace of the study.

When the participant arrived, they were greeted and guided to the quiet room. Instructions were provided to inform them about the tasks in chronological order. A practice trial was presented for them to get used to the task and reduce the probability of errors occurring. Afterwards, they were asked to sit comfortably and were positioned at the appropriate distance and angle to the eye tracker, using the corresponding Tobii Pro Lab tool. Participants were seated at approximately 65 cm from the screen and were free to move their heads during data collection, as no head stabilization system was used. Nonetheless, before the reading episode, participants were instructed to minimize head movements as much as possible. Once adequately positioned, a 9-point calibration followed by a subsequent 4-point validation was conducted. This process was repeated until satisfactory parameters were achieved. However, it was also noted that increasing the number of calibration or validation trials could lead to participant fatigue, potentially causing accuracy, precision, and data loss parameters to exceed the recommended thresholds ($\leq 0.5^\circ$ for accuracy, $\leq 0.2^\circ$ for precision, and $\leq 10\%$ for data loss).

After calibration and validation, the data recording started, lasting a maximum of 10 min. First, the participant was required to read the instructions (*"Please read the following texts on the homework debate carefully in order to answer the following question afterwards: Do you think assigning homework is a useful teaching method? Please support your conclusions based on the arguments expressed in the texts"*). Following this, the four texts were presented. The participant determined when to proceed to the next screen by clicking the mouse they had

at hand, understanding that each text occupied two sequential screens and there was no option to revisit the screen once advanced. Each fragment of each text was positioned in the center of the screen, leaving a 15% margin (Figure 3). Furthermore, based on previous studies (e.g., [62]), the text was written in black Courier New 35 font on a light grey background and presented in double interlinear spacing. In addition, the presentation resolution was 2560 x 1600 pixels and the scale 100%. The screen brightness was also adjusted to enable comfortable reading. Before each screen, a cross (X) was presented for 800 milliseconds. The cross was located to the left of the first word to guide where the first fixation of the stimulus should occur. Furthermore, as commented, to avoid potential interference caused by the order of presentation of the texts, they were presented randomly to each participant using the shuffle option in Tobii Pro Lab. Following the reading episode, they were instructed to write the essay, which took about 30 min. During the reading and writing intervals, the experimenter was seated in such a way that the participant could be observed but the participant could not see the experimenter. Finally, the participant was thanked and debriefed.

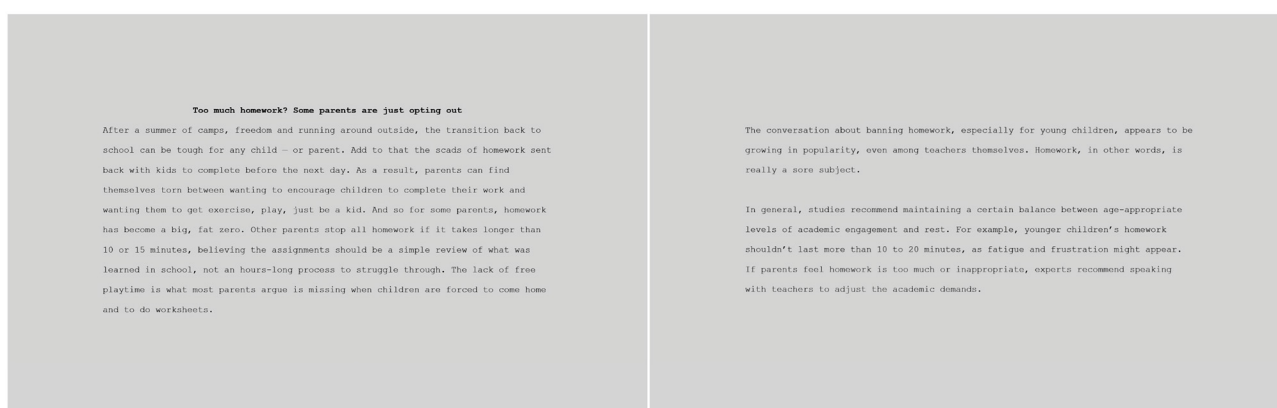


Figure 3. Example of a text displayed on two screens in Tobii Pro Lab.

2.6. Data Analysis

Once the eye movements had been recorded, data cleaning procedures were conducted to ensure the use of fixations that faithfully reflected lexical processing [63]. First, the Tobii I-VT fixation filter was selected, as it is appropriate for controlled studies with minimal head movement in which the collected data contain only fixations or saccades [64]. This means that Tobii Pro Lab automatically eliminated all fixations lasting less than 60 msec and merged fixations separated by a maximum time of 75 ms and a maximum angle of 0.5°. Secondly, decisions were made about whether to include, partially include, or exclude each participant from analyses involving gaze measures. These decisions were based on a review of the eye movement recording of each participant. For instance, considerations included whether the participant's gaze followed a natural flow, landed on the sentences rather than on the blank areas, etc. Thirdly, extreme values were manually eliminated for each participant across the twenty arguments. This procedure separated arguments in favor from arguments against and then calculated the mean and standard deviation for each participant for first-pass reading duration on arguments in favor, first-pass reading duration on arguments against, lookback reading duration on arguments in favor, and lookback reading duration on arguments against. Any values falling beyond three standard deviations from the mean were eliminated [63]. Subsequently, the four aggregated eye movement dependent variables were calculated for each participant.

Regarding the statistical analyses, IBM SPSS Statistics software (version 25.0) was used to conduct all of them. Decisions were made regarding which analyses to run in accordance with the study design and the assumptions considered for each potentially suitable statistical test. For the first research question, two repeated measures analyses of covariance (ANCOVA) were conducted to examine the relationship between prior beliefs and first-pass

reading duration and the relationship between prior beliefs and lookback reading duration. In both ANCOVAs, prior beliefs were included as a covariate, and a within-subjects factor was created to reflect the manipulation of the argument stance: arguments in favor vs. arguments against. Although the variables did not fully meet all the assumptions (e.g., normality or linearity), interaction effects between the covariate and the within-subjects factor were explored. Interaction effects capture how the relationship between variables changes across different levels of another variable, and this differential relationship can provide meaningful insights even when some assumptions are not strictly met [65]. However, as results should be interpreted cautiously, an additional non-parametric test was conducted: Spearman's rank correlation. In relation to the second research question, which assessed the relationship between prior beliefs and essay's positioning, a regression analysis was initially considered, but due to violations of assumptions again, Spearman's rank correlation was employed instead.

3. Results

3.1. Descriptive Statistics for Prior Beliefs

As mentioned in the previous section, a five-point single-belief score ranging from 1 (*completely against*) to 5 (*completely in favor*) was computed from participants' responses to six items (five after removal of item b). Overall, participants demonstrated a slight preference for assigning homework as a teaching method ($M = 3.42$, $SD = 0.75$, range: 2.2–4.6).

3.2. Research Question 1: Prior Beliefs and Eye Movements

After removing completely two participants from the study due to a generalized poor registration of their eye movements, adequate data quality parameters were obtained for the remaining participants (accuracy: $M = 0.40$, $SD = 0.19$; precision: $M = 0.19$, $SD = 0.07$; data loss percentage after the calibration-validation procedure: $M = 0.8$, $SD = 1.5$; data loss percentage after the reading episode: $M = 5.77$, $SD = 2.86$). Despite this, four participants had some of their eye tracking measurements removed because of a poor calibration only in certain sections of the screen (e.g., only on the upper half of the screen but not on the lower half). Consequently, out of the total of twenty AOIs identified across the four texts, these four participants were left with 10, 11, 13 and 16 AOIs each. Regarding the outlier removal, no data were eliminated. Table 4 shows the final descriptive statistics for the four gaze measures.

Table 4. Descriptive statistics for the gaze measures.

Variables	Minimum	Maximum	Mean	SD
Fp_ArF	107.29	371.44	214.46	73.34
Fp_ArA	117.08	326.46	185.61	57.26
Lb_ArF	74.88	533.59	240.61	128.08
Lb_ArA	34.97	394.54	180.73	110.97

First-pass reading duration for arguments in favor (Fp_ArF); first-pass reading duration for arguments against (Fp_ArA); lookback reading duration for arguments in favor (Lb_ArF); lookback reading duration for arguments against (Lb_ArA). The unit for the four variables is milliseconds/word.

Regarding the two repeated measures ANCOVAs to address the first research question, contrary to expected, the covariate (prior beliefs) was not found significantly related to the within-subjects factor (argument stance: in favor vs. against), neither for first-pass reading duration, $F(1, 15) = 0.017$, $p > 0.05$, $\text{partial } \eta^2 = 0.001$, nor for lookback reading duration, $F(1, 15) = 0.094$, $p > 0.05$, $\text{partial } \eta^2 = 0.006$. These results are coherent with the Spearman's rank correlations obtained, since as it can be observed in Table 5, not much difference was found in the correlations between prior beliefs and gaze measures depending on the argument stance. Furthermore, only one of these four correlations was found to be significant, which is the relationship between prior beliefs and first-pass reading duration for against arguments ($r_s = 0.501$, $p < 0.05$).

Table 5. Spearman’s rank correlations between prior beliefs and each gaze measure.

Gaze Measures	Prior Beliefs	
	r_s	Sig.
Fp_ArF	0.471	0.056
Fp_ArA	0.501 *	0.041
Lb_ArF	−0.171	0.512
Lb_ArA	−0.021	0.937

* $p < 0.05$.

3.3. Research Question 2: Prior Beliefs and Essay Performance

As commented, three measures were obtained from the participants’ performance on the essay: sum of arguments in favor, sum of arguments against, and a positioning score, i.e., the difference between the two aforementioned indices, ranging from −10 (*completely against*) to 10 (*completely in favor*). On average, as reflected by the positioning score, participants included a similar number of arguments in favor and arguments against ($M = -0.35$, $SD = 2.35$, range: −5–5). To approach the second research question, i.e., the relationship between prior beliefs and the essay’s positioning, a Spearman’s rank correlation was conducted. In this case again, contrary to expected, no significant relationship was found between these two variables ($r_s = 0.225$, $p > 0.05$).

4. Discussion

4.1. Summary of Evidence

The goal of the present study was to investigate the extent to which recent upper secondary graduates’ prior beliefs on a controversial issue are related to their behavior while reading (eye movements) and to their positioning in a written essay elaborated by them after reading the conflicting texts. Based on previous international surveys, such as PISA [66] and PIAAC [4], there are literacy deficiencies among students completing basic compulsory education and young adults who have recently finished upper secondary education. Specifically, according to research on reading comprehension (e.g., [10,11]), these deficiencies manifest when reading multiple contradictory documents. Both adolescents and undergraduates tend to allocate more cognitive resources to belief-consistent arguments, leading to an unbalanced and one-sided mental representation of the texts. This phenomenon is known as the text-belief consistency effect (e.g., [7–9]), which was the focus of the present study. Although the data obtained in this study provide insight into the research questions, the results do not generally support previous research findings [16,17] or the proposed hypotheses derived from them, as will be now explained.

First, we expected to find a longer first-pass reading duration for belief-inconsistent arguments (Hypothesis 1), indicating that participants have detected a discrepancy between their prior beliefs and the text information. This detection of a discrepancy would suggest that the readers were passively validating the text information based on their prior beliefs, which aligns with the first step proposed by the two-step model of validation [10], known as routine validation. The results obtained through the first repeated measures ANCOVA indicate that the data do not provide sufficient evidence to confirm that prior beliefs differentially affect the time spent reading an argument for the first time depending on its stance (in favor or against). This would suggest that belief-inconsistent arguments have not necessarily provoked a disruption in comprehension, contrary to what would be expected during routine validation for all readers, regardless of whether they attempt to resolve the discrepancy or not [10]. However, the Spearman’s rank correlation between prior beliefs and first-pass reading duration for arguments against was found to be positive and significant. This indicates that the more participants agreed with using homework as a teaching method, the more time they spent reading arguments against during the first pass, which is consistent with Hypothesis 1. Despite this, the Spearman’s rank correlation between prior beliefs and first-pass reading duration for arguments in favor was not found to be significant. Not only that, but there is a positive relationship between these two

variables, contrary to what would be expected (i.e., the more participants disagreed with using homework as a teaching method, the more time they should have spent reading arguments in favor during the first pass). This unexpected relationship can explain why the repeated measures ANCOVA did not find a significant interaction effect between prior beliefs and the within-subjects factor. Therefore, Hypothesis 1 cannot be completely confirmed and does not entirely align with prior evidence employing this same metric (first-pass reading duration) to explore routine validation [17]. However, there is some evidence in the expected direction.

Second, as commented, detecting a discrepancy does not necessarily mean that the subject tries to resolve the conflict. Therefore, in line with the text-belief consistency effect, we expected to observe a longer lookback reading duration for belief-consistent information (Hypothesis 2), indicating that participants did not engage in the second step of the two-step model of validation [10], namely, the strategic validation of the belief-inconsistent information. The results obtained through the second repeated measures ANCOVA indicate that the data cannot provide sufficient evidence to confirm that prior beliefs differentially affect the time spent reading an argument during lookbacks depending on its stance (in favor or against). In this case, no significant Spearman's rank correlation was observed. These findings imply that Hypothesis 2 cannot be confirmed. Therefore, participants' reading patterns could suggest that they have overcome the text-belief consistency effect. Since the results do not indicate a preference for deeply processing either belief-consistent or belief-inconsistent arguments, this suggests that readers are allocating the same amount of cognitive resources to both types of arguments. This equivalent reinspection behavior is contrary to what would be expected if the text-belief consistency effect had not been overcome, which is precisely what prior eye tracking studies on this effect suggest occurs [16,17].

Similarly, we anticipated finding an unbalanced inclusion of arguments in their written essay, with a preference for belief-consistent arguments (Hypothesis 3), reflecting a biased mental representation of the text content. However, the Spearman's rank correlation has not been found significant, and thus, Hypothesis 3 cannot be confirmed either. This lack of relationship between these two variables could suggest that readers are incorporating a similar amount of belief-consistent and belief-inconsistent arguments in their essays, once again implying that participants may have overcome the text-belief consistency effect and, consequently, have engaged in the strategic validation step of the two-step model of validation [10]. According to the descriptive statistics, the essays' average positioning score was -0.35 , with a standard deviation of 2.35 and a range between -5 and 5 . Considering these values, and the potential range of -10 to 10 for the positioning score, it seems plausible to conclude that readers exhibited a balanced consideration of arguments regardless of their prior beliefs. However, once again, these results do not align with prior evidence exploring the text-belief consistency effect, in this case, using essays as an assessment method [16,31–34].

Therefore, results might be indicating that recent upper secondary graduates can overcome the text-belief consistency effect, which is observed not only in their behavior while reading (eye movements), but also in their written products after reading. This finding could mean that recent upper secondary graduates have learned to manage multiple contradictory documents by associating this reading context to the need to employ deep processing strategies (i.e., strategic validation, in line with the two-step model of validation [10]). Therefore, according to the RESOLV model [2,19], when these readers encounter multiple contradictory texts, this association would be activated in the readers' context model, and thus, it would influence the construction of their task model, that is, the mental representation of the goal to be fulfilled and the means to achieve it. These unexpected positive results may be due to the effectiveness of the current secondary and upper secondary school curriculum in Spain, which incorporates competences related to the comprehension of multiple sources, including contradictory ones. These curricula were implemented for the first time at the end of 2022, well after the last PIAAC assessment (2011–2012) and not arriving in time for the most recently published PISA assessment (at

the beginning of 2022). To explore this possibility, further research should examine the development of these skills across various ages and educational levels. Additionally, classroom practices should be observed and analyzed to understand how these competences are being taught and assessed, establishing comparisons between schools and across different regions of Spain that follow distinct guidelines. Furthermore, it would be interesting to explore differences in reading behavior and performance not only within Spain but also between different countries. In this light, a degree of coherence with the upcoming PISA and PIAAC results should be anticipated.

Nevertheless, some aspects need to be considered before generalizing this statement to a wider population. The present study's sample was small ($n = 17$) and restricted to a particular gender (women) and to a particular situation (undergraduate students enrolled in the first year of the bachelor's degree in Primary Education Teacher Training at a Spanish university). Sample size issues are common among eye tracking studies. Specifically, as has been observed, data attrition due to poor recording quality is often a significant challenge (e.g., [17,52,67]). Moreover, as will be discussed later, thorough cleaning methods should be applied to the remaining sample to eliminate any errors or behaviors that do not accurately reflect the processes being studied. Therefore, if the sample size is large, the data cleaning procedures can become extremely time-consuming; thus, a smaller sample size may be more feasible. However, a small sample size may lead researchers to commit a type II error, meaning they dismiss results as null when, in fact, an effect is present. For these reasons, as Holmqvist et al. indicate [13], to make studies with small sample sizes valuable, they should not be presented as fully generalizable experiments, and hypotheses should be generated for later testing.

Furthermore, it is worth mentioning that the students that volunteered to participate could be particularly engaged in the task compared to those students in the same class who did not volunteer. According to Richter and Maier [10], motivation is one of the conditions that contribute to deeply processing text information and thus to overcoming the text-belief consistency effect. Moreover, results may also result from the fact that the sample consists solely of one type of recent upper secondary graduates, first-year undergraduates, who may have particularly acquired the skills. Including a more diverse sample, beyond those who have entered university, might yield different results. Additionally, the instruction given to the participants before the reading episode could have also been a contextual factor influencing the construction of their reading goal and consequently affecting their reading behavior and performance. As Stadtler et al. [1] found in their research comparing the instruction to write an argument, a summary, or a list of keywords, only participants in the first condition included a two-sided approach to the debate in their written essay. For all these reasons, the statement that recent upper secondary graduates can overcome the text-belief consistency effect should be interpreted with caution.

Despite all of this, and returning to Hypothesis 1, if the results obtained from the first repeated measures ANCOVA were to be interpreted as a complete absence of discrepancy detection by participants between their prior beliefs and the text information, it would be necessary to revise the interpretation of the subsequent hypotheses. According to the two-step model of validation [10], routine validation is an unconscious and passive process expected to occur when individuals read controversial topics. However, this can only happen if prior beliefs have been sufficiently activated by contextual cues (e.g., by concepts or propositions participants encounter in the texts while reading). Therefore, the observed absence of a differentiated duration in first-pass fixations between arguments in favor and arguments against could be indicating that participants' prior beliefs were not activated in their working memory while reading. If beliefs were not activated, they could not have been used for the involuntary validation of text information, and incoherences could not appear within readers' mental representations. Therefore, it would not make sense to distinguish between belief-consistent and belief-inconsistent texts or arguments. If this interpretation is correct, it cannot be guaranteed that valid data have been gathered regarding the relationship between prior beliefs and lookback reading duration (Hypothesis 2) and

between prior beliefs and essay's positioning (Hypothesis 3). Consequently, there would not be enough evidence to determine whether recent upper secondary graduates overcome the text-belief consistency effect or not.

However, it is unlikely that prior beliefs were not activated, as the texts used in the present study were also employed in previous research, where a text-belief consistency effect was detected [58]. Furthermore, although it could have been possible to include questions just before the reading episode to activate participants' beliefs, this is not usually done. Researchers try to minimize carry-over effects, that is, the potential influences that an earlier part of a study might have on later parts (e.g., [11,16]). Despite not including these initial questions, authors were also able to detect a text-belief consistency effect. Hence, if these results cannot be explained by the lack of activation of prior beliefs, they could be attributed to the sample's average prior positioning being close to neutral. Research suggests that disruption in comprehension is greater when reader's prior beliefs are stronger (e.g., [16,56]). Since participants showed only a slight preference for assigning homework as a teaching method, and their scores did not fluctuate much, this explanation is plausible. Nevertheless, the absence of differences in the first-pass reading time between arguments in favor and arguments against may also be due to the metric's potential insensitivity to inconsistencies, in line with the findings of Maier et al. [16] and Rinck et al. [40]. Different results might have been observed if first-pass re-reading times had been analyzed; however, Tobii Pro Lab does not offer this metric.

4.2. Limitations

Despite having found interesting results, the present study has some limitations that need to be addressed. First, although eye tracking methodology offers several benefits, such as providing direct and objective data, even from the most unconscious cognitive processes, this technology also presents certain difficulties that could have impacted the research. On the one hand, although the average data quality results are more than acceptable (i.e., all of them fall within the established thresholds), these values did not ensure a complete accurate and precise recording of eye movements. That is why two participants were entirely removed from the analyses, and four others were reviewed to include only those AOIs where the eye movement registration could be considered reliable. While the other thirteen participants were not thoroughly reviewed, there is a high probability that a similar, albeit much less significant, degree of inexactness was present in these participants as well. These issues arise due to the decision to include texts as the stimuli to be processed, within which sentences were selected to be the AOIs, instead of simply presenting sentences or words separately. Consequently, the AOIs were relatively small and closely spaced, making inaccuracies and imprecisions more likely to occur. However, the stimuli employed are determined by the goal of the study, namely, to explore the comprehension of multiple texts, and thus, this aforementioned inexactness is an inevitable challenge to be faced in this area of study. A solution to obtain higher quality data could be to make the accuracy and precision thresholds even stricter. However, it would not be viable as it would lead to endless calibration-validation cycles to obtain the desired quality results or to a high exclusion rate of participants.

On the other hand, it seems that there could be certain overlapping between the two indices of visual behavior registered in this study: first-pass reading duration and lookback reading duration. For instance, if during the first pass the reader exits a sentence and then goes back to reading the unread parts of this same sentence, these new fixations are registered as a lookback, even if the participant is reading these parts of the sentence for the first time. Rather than a reinspection of the sentence, the reader experienced an interruption and is now resuming reading it. Thus, interpreting lookback reading duration as the amount of reinspection would not be accurate in this case. From the opposite point of view, while participants are reading a text, their attention may be captured by a salient word in another part of the text, although they may not begin to read that sentence for comprehension until some time later. Therefore, the real first pass could be

codified as a lookback, whereas that first fixation, which was just a bottom-up process, could be codified as the only first-pass fixation. This last phenomenon raises the question of whether the interpretation given to a certain gaze measure is correct, as there could be other cognitive processes happening underneath that manifest through the same eye movements (e.g., first-pass duration may indicate both bottom-up attention processes and top-down comprehension processes, depending on the context). Typically, these aspects are controlled by offering text material designed a priori to minimize irrelevant and disrupting processes [41], e.g., removing potential distractors, complex vocabulary, etc. However, this issue is increasingly motivating researchers to seek ways to validate eye tracking indicators [68].

Further efforts to master the eye tracking device (Tobii Pro Nano) and software (Tobii Pro Lab) need to be made to comprehend how to overcome these issues. However, a clear but effortful solution that could be applied would be to clean gaze data even more exhaustively. Tobii Pro Lab can provide a database including every single fixation, their duration, and the AOI in which they have been registered. By comparing this database with the participants' recording of their eye movements, more detailed cleaning procedures can be implemented. Not only that, but this procedure would also allow researchers to obtain other gaze measures that Tobii Pro Lab does not offer, such as first-pass re-reading duration, lookback count, or integration measures. However, as commented, it is a complex and time-consuming endeavor.

Other complementary strategies to mitigate these issues are as follows. First, fixations can be merged or removed according to evidence-based practices [63], such as those applied in this study: merging fixations separated by a maximum time of 75 msec and a maximum angle of 0.5° and removing all fixations lasting less than 60 ms. Additional practices not applied in this study include removing longer fixations, either when they exceed a general threshold or a participant-specific threshold. Furthermore, recent methods have been developed to correct vertical drift (i.e., the displacement of fixation registrations along the vertical axis), e.g., Carr et al.'s method [69]. Finally, as noted in the method section, errors can be minimized by optimizing conditions, such as providing adequate lighting, ensuring compatible glasses, or removing potential distractors [13]. We aimed to create optimal conditions as much as possible, though we encountered some unexpected issues (e.g., participant tiredness, outdated glasses, or contact lenses).

Second, there are some aspects of the other measures and the reading material worth mentioning. Regarding the scale used to measure prior beliefs, as commented, the items were extracted from a questionnaire utilized in a previous study [58]. This questionnaire was specifically created by the authors to suit the context of the dispute presented in the controversial texts, that is, homework as a teaching method. Consequently, this scale has not been widely used and tested, unlike Maier and Richter's belief scale [7]. However, Maier and Richter's scale was specifically designed to measure beliefs about global warming and vaccination. Changing the theme of the texts from homework to either of those two would have somewhat reduced the ecological validity, as the texts used were similar to the actual learning material employed by the study's sample, i.e., preservice teachers [15]. Furthermore, the language of the texts is an important factor to be considered. As research has revealed (e.g., [70]), L2 (second language) readers may have additional challenges when reading, as their cognitive resources may be consumed by lower-level processes (e.g., word decoding, syntactic processing, etc.) at the cost of higher-level processes (e.g., multiple document integration, evaluation of the sources, content and form, etc.). However, all the participants belonged to an English teaching group, where a proficiency corresponding to B2, according to the Common European Framework of Reference for Languages, is expected. Despite this, their English proficiency could have been incorporated as a covariate to explore or control its effects. Another alternative could have been using participants' first language (L1) for the texts.

Finally, the identification of arguments within the texts and within the participants' written essays could have benefited from inter-rater agreement. Additionally, instead of classifying sentences within texts as either arguments or irrelevant information, it might have been advantageous to introduce a more detailed classification for arguments. Arguments can be categorized as a claim, i.e., a statement that the author wants the reader to accept as true, or as a reason, i.e., data or information that supports or justifies the claim [71]. Prior research has indicated that beliefs appear to play a critical role in the processing of claims, yet they tend to be less significant when considering reasons [16,17,56]. Regarding the scoring of the essays, a broader framework of arguments in favor or against could have been employed, not restricted to those included in the four texts. Participants often included arguments which had not been presented in the texts, and strictly speaking, these arguments should have not been considered. However, the exclusion of these arguments would have biased their positioning score and thus the study results.

4.3. Gaps of Knowledge and Practical Implications

Results from the present study could be indicating that recent upper secondary graduates effectively overcome the text-belief consistency effect. However, due to the various limitations identified in the study, further research should be conducted to confirm these findings. In these new studies, a bigger and more representative sample should be used, ideally employing an a priori power analysis to estimate the necessary sample size and ensure equal representation of different subpopulations. Moreover, some variables that could be controlled or explored in further research, besides prior beliefs, could be participants' age, gender, study discipline, prior educational studies, background knowledge of the text theme, reading skills, L2 proficiency, etc. In the present study, the participants' degree access score was initially planned to be included as a covariate. However, statistical analyses revealed that this variable was not correlated with any variables, and thus, it was removed from the analyses. These results may be attributed to the small and homogeneous sample size. Therefore, as previously noted, increasing both the size and diversity of the sample is a crucial consideration for future research. In these cases, a more sophisticated analysis, such as multilevel modeling, could prove beneficial in addressing the role of individual differences in the context of the text-belief consistency effect.

Furthermore, despite controlling the order of presentation of the texts by displaying them randomly, this may not have been sufficient due to the small sample size as well. As many studies have shown (e.g., [7,9,16]), presenting texts in an alternating manner (e.g., first presenting a belief-consistent document, followed by a belief-inconsistent document before returning to a belief-consistent text) can lead to deeper processing compared to presenting texts in a blocked manner (i.e., individuals reading documents of the same belief type successively). Therefore, taking Abendroth and Richter's study as an example [17], to control for the effect of presentation order, future studies could employ a blocked presentation format for all the participants.

In addition to all this, more rigorous cleaning procedures for gaze data, as well as a more detailed classification of text arguments, and a broader framework for scoring written essays should be considered by further research. Moreover, to achieve a comprehensive and valid depiction of students' processes and performance, additional assessment tools should be incorporated, especially considering the limitations of the two measures employed, which have been extensively discussed previously. Qualitative methods, such as interviews or think-aloud protocols, could be particularly beneficial for exploring the processes occurring during reading, allowing for a comparison between these methods and eye tracking technology. However, it must also be acknowledged that interviews and think-aloud protocols have significant limitations that are especially relevant to the present study, e.g., difficulties in accessing unconscious processes and biases related to self-report. Since each method has its own advantages and disadvantages, combining different instruments can yield rich data and fruitful findings, making it worthwhile despite the time-consuming nature of the process.

Further investigations using eye tracking methodology could explore the text-belief consistency effect in different populations, especially children ending basic compulsory education (around 15 years old). Eye tracking studies exploring this effect have focused on university students [16,17]; however, the end of basic compulsory education is a crucial point in educational systems. As education will no longer be compulsory from this point onwards, students are expected to have acquired the competences that will prepare them for further education and adult life. Moreover, future studies, whether focusing on secondary education students or university students, could explore additional variables, such as the establishment of inferences or attention to sources, and investigate how these variables are influenced by the stance of the text and the reader's prior beliefs.

The findings obtained by these future investigations could also prove very useful in refining or providing additional evidence for theoretical frameworks, such as the PISA 2018 reading framework of processes [21], or more recent proposals, such as the Proficient Academic Reader (PAR) framework [72–74]. For instance, PAR has been specifically used to study the essential reading skills for university students. These skills encompass foundational reading abilities (e.g., word-level or sentence-level processes), as well as strategic reading, task awareness, and motivation to advance through the task. PAR advocates for evaluating these skills using situated measures, that is, assessing student performance while they are engaged in the task, instead of relying on retrospective or offline reports, which are often inaccurate [75]. Situated measures can either be behaviorally grounded, involving the quantification of a participant's behavior while reading, or be based on retrospective judgments made immediately following the reading episode. Therefore, eye tracking methodology aligns well with this purpose.

However, to not only employ situated measures but also to establish an ecologically valid context during reading, efforts must be made to devise feasible and secure methods for integrating eye tracking devices into real-life learning environments. When conducting studies in laboratory settings, where one participant at a time is tested with minimal disturbance, an environment very different from educational practice is created [15]. An alternative that could guarantee ecological validity is the use of glasses integrating the eye tracking system. For instance, Tobii, the same company that developed and manufactured Tobii Pro Nano and Tobii Pro Lab, also commercializes Tobii Pro Glasses 3, which have been designed to work in various environment, from cars to classrooms [76]. In addition to this, to further enhance ecological validity, it is important to ensure that the task is meaningful for the participants, i.e., it mirrors an actual learning or reading situation for them. An example already discussed is the theme of the texts, but it could also include the text layout, the type of instructions and contextual cues provided, or the type of responses requested from them. The concept of meaningful tasks is not new. Wiggins' authentic assessment referred precisely to using realistic tasks that replicate or simulate real-life contexts [77].

With respect to other implications for educational practice, as mentioned in the introduction, obtaining a deep understanding of how students read multiple contradictory documents can help educational stakeholders better direct their efforts to promote the skills needed to master these reading situations. Despite the present study indicating that recent upper secondary graduates may be capable of overcoming the text-belief consistency effect, it is important to consider the identified limitations and the evidence from prior studies suggesting otherwise. Hence, if we assume that there is a widespread deficiency among students, measures need to be implemented at schools, and these measures must take into account the findings obtained by this prior research. First of all, curricula serve as a point of reference for teachers. A generalized issue among students in achieving a specific competence may indicate that curricula are not effective—either because they do not adequately address the competence or because they are written in a way that makes it difficult for teachers to translate them into their day-to-day practices. Therefore, governments may need to change their policies. To do this, they should consider research that explores different curricula (e.g., as commented, comparing various regions within the

same country or examining differences between countries) and how these variations can be related to differences in the acquisition of the corresponding competences.

Secondly, schools could participate in programs designed to acquire and develop higher-level reading skills. One example is self-explanation reading training (SERT) [78,79], a human-delivered program that has proven successful for both secondary school [80] and university students [78]. This program also served as the basis for developing the interactive strategy trainer for active reading and thinking (iSTART) [81], an automated and interactive training program that has similarly been shown to benefit secondary school [82] and university students [83]. Both programs target comprehension monitoring (i.e., recognizing a failure in understanding and subsequently using additional active strategies), along with other reading processes such as paraphrasing, establishing bridging inferences, and making predictions and elaborations. In addition to these programs, teachers could create meaningful situations where students can learn to apply these higher-level reading skills to real-life and practical contexts. These tasks can be proposed from an interdisciplinary approach, that is, converging different subjects to make the task even more meaningful for the students. For instance, a potential activity could involve the designing of a digital multilingual newspaper. To complete this task, students would need to read multiple documents on the same topic from different, and possibly contradictory, sources to achieve an integrated understanding of the issue.

Furthermore, these types of tasks could be accompanied by aids to guide students' cognitive processes. For example, when reading multiple discrepant texts, teachers' feedback could direct students to identify their prior positioning on the debate, detect inconsistencies, reread the texts, and focus on those arguments that are incongruent with their prior beliefs. This guidance would only be possible thanks to the eye tracking studies that have identified the cognitive processes underlying the text-belief consistency effect. In addition to this, within the research domain of eye tracking and reading comprehension, there exists another tool that could fulfill this guidance function: eye movement modeling examples (EMME). EMMEs are rooted in observational learning [84] as they present video-based examples of eye movements performed by experts during a specific task. EMMEs can enhance processing in those who observe them by directing their attention. Specifically, learners synchronize their visual attention with the model's, thus engaging in joint attention [85]. There is ample evidence demonstrating the effectiveness of EMMEs in reading comprehension [62,86–88], including critical reading of multiple conflicting sources [51]. Therefore, teacher's feedback could be enhanced by the additional use of EMMEs.

Importantly, if it is confirmed that recent upper secondary graduates are generally able to overcome the text-belief consistency effect, the aforementioned practices (meaningful tasks, EMME, and feedback) could still be beneficial, particularly for those students who do not meet the standards. These measures could be supported by eye tracking studies such as the present one, as they help identify the eye movements associated with specific cognitive processes involved in effective reading comprehension.

Finally, eye tracking research can also be used to design materials that better support reading comprehension [89]. For example, researchers can manipulate the font size, color, and layout of a text to observe in which cases eye movements correspond to the desired cognitive processes. Furthermore, by identifying which parts of a text attract readers' attention the most, how much time readers spend on them, and whether they regress to previous sections or not, instructional materials can be optimized. Therefore, eye tracking methodology has great potential for educational practice and, once its limitations are effectively addressed, will offer significant benefits for acquiring literacy proficiency.

5. Conclusions

Nowadays, due to increasing exposure to information from the World Wide Web, individuals face the difficulty of making decisions and completing diverse tasks based on multiple and conflicting sources of information. For individuals to be autonomous and responsible citizens who do not succumb to manipulation and misinformation, schools

must work on ensuring that students are prepared for this challenge. According to international surveys, such as PISA [66] and PIAAC [4], both adolescents and young adults show deficiencies in their literacy skills. In particular, research on reading comprehension (e.g., [10,11,16,17]) shows that students generally tend to favor information consistent with their prior beliefs when processing texts presenting discrepant views, a phenomenon known as the text-belief consistency effect [7–9].

The present study has gathered gaze and performance data to determine whether recent upper secondary graduates overcome the text-belief consistency effect. Results suggest that this population may be capable of overcoming this bias, as deep processing, indicated by lookback reading duration, does not differ depending on whether the arguments are belief-consistent or belief-inconsistent. Moreover, participants' written essays show a balanced inclusion of both types of arguments. However, the study presents some limitations that urge caution in making interpretations. Despite these issues, the present research contributes to the body of studies utilizing eye tracking methodology in the reading of multiple texts, particularly those examining the text-belief consistency effect, which are currently scarce, likely due to its inherent complexity. Not only that, the gaze measures are complemented with performance data. In light of the results, this study proposes hypotheses for future exploration while elucidating the utility of eye tracking methodology for research in this area. We encourage further studies, to the extent possible, to employ larger and more representative samples, explore individual differences, and include a mixed-method approach, while also implementing more exhaustive cleaning procedures for the gaze data. Only then will generalizable and valid conclusions be possible, which can be used by educational stakeholders to enhance these skills, whether among the general student population or within specific groups who struggle to meet the standards.

Author Contributions: Conceptualization, M.G.-S., I.M. and R.C.; methodology, M.G.-S., I.M. and R.C.; validation, M.G.-S. and I.M.; formal analysis, M.G.-S. and I.M.; investigation, M.G.-S.; resources, I.M. and R.C.; data curation, M.G.-S.; writing—original draft preparation, M.G.-S.; writing—review and editing, M.G.-S., I.M. and R.C.; visualization, M.G.-S.; supervision, I.M.; funding acquisition, R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by Ministerio de Ciencia, Innovación y Universidades, under Plan Estatal de Investigación Científica y Técnica y de Innovación 2017–2020 (grant number PID2020-117996RB-I00).

Institutional Review Board Statement: Ethical review and approval were waived for this study because it was conducted with adults and did not involve the processing of sensitive personal data as defined in Article 9.1 of the EU General Data Protection Regulation. Specifically, the data collected in this study do not concern an individual's racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership status, health or sex life, or sexual orientation. Additionally, approval from the Committee of Ethics and Human Research at the University of Valencia was not required, as this study is part of a master's thesis, which receives approval from the academic supervisors. Furthermore, all participants were fully informed about the research and provided their informed consent prior to participation.

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

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Conflicts of Interest: The authors declare no 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.

Appendix A Text Material

After a summer of camps, freedom and running around outside, the transition back to school can be tough for any child — or parent. Add to that the scads of homework sent back with kids to complete before the next day. As a result, parents can find themselves torn between wanting to encourage children to complete their work and wanting them to get exercise, play, just be a kid. And so for some parents, homework has become a big, fat zero. Other parents stop all homework if it takes longer than 10 or 15 minutes, believing the assignments should be a simple review of what was learned in school, not an hours-long process to struggle through. The lack of free playtime is what most parents argue is missing when children are forced to come home and to do worksheets.

The conversation about banning homework, especially for young children, appears to be growing in popularity, even among teachers themselves. Homework, in other words, is really a sore subject.

In general, studies recommend maintaining a certain balance between age-appropriate levels of academic engagement and rest. For example, younger children's homework shouldn't last more than 10 to 20 minutes, as fatigue and frustration might appear. If parents feel homework is too much or inappropriate, experts recommend speaking with teachers to adjust the academic demands.

Figure A1. Text “Too Much Homework? Some Parents Are Just Opting Out”.

Homework provides an opportunity for parents to interact with and understand the content their children are learning so they can provide another means of academic support for them.

Furthermore, research has identified perseverance as an essential skill children develop when they complete homework. Most students will not manage to solve all the tasks the first time without guidance from their teacher. They need to try again and again, persevering until they succeed. Also, research has proven that regular accomplishments like finishing homework build self-esteem. Being able to consistently submit the work on time and do a great job at it immediately makes them feel confident about what they are capable of.

Homework does indeed serve a purpose, with studies providing a link between homework and higher secondary school attainment. However, benefits are not always as apparent as in these cases, leading to disagreements about the need to ask for assignments to the younger students. Despite of this, homework helps students in primary school and younger years to prepare them for secondary school, encouraging them to become independent learners.

Although all of this is true, it must be taken into consideration that not just any type of task will do. Homework being set should have a clear goal and purpose, so as to encourage student completion of the work and for it to be truly worthwhile.

Figure A2. Text “Homework: Supporting Student Success and Growth”.

A primary school teacher from an urban public school has decided that she will not be assigning homework to her students this year. She wants them to set aside the books when class is over. Beatriz García, the elementary school teacher, has passed out a letter to parents after the summer break. “After much thinking this summer, I am trying something new”, she wrote in the letter. “Research has proven that homework can hinder student learning and performance, as some of its negative effects include decreasing students’ motivation and interest”.

“Rather, I ask that you spend your evenings doing things that are proven to correlate with student success: spend time as a family, play outside and get your child to bed early”, she adds.

The debate over whether to assign homework has been ongoing for some time, sparking interest among teachers, parents, researchers, policymakers, and educational stakeholders alike. According to research from the Ministry of Education, the age of the students, the type of assignments and the amount of study time are important factors to take into consideration. “The central lesson of this body of research is that homework is not a strategy that works for all children”, the group said.

They also suggest that “heavy homework loads should not be used as a main strategy for improving home-school relations or student achievement”.

Figure A3. Text “A Primary School Teacher Stopped Assigning Homework”.

Homework has been a part of the class ever since schools were established and “do my homework” is the most common question among students. Teachers assign homework so that pupils have something to work on in their spare time. There is nothing bad about this idea. Here are some of the most important homework benefits:

- One of the pros of homework is the fact that it teaches students how to stay organized, plan ahead and break complex tasks into smaller, more manageable parts. When one has to complete 4 different assignments in 3 days, one tends to be a lot more organized and attentive. As a consequence, homework makes students more responsible.
- Of course, homework gives students another chance to review class material. In many cases, students are not paying enough attention during class to learn everything. However, when they receive homework, they have to reread everything they wrote down (and even find additional information about some topics) to complete the assignments.
- Another benefit is the fact that it helps the teacher assess how much of the class was understood by the students, offering insights into aspects where students may need further clarification. This allows the teacher to tailor his or her teaching methods and adjust the lesson plans accordingly, ensuring that every student has the opportunity to fully understand the subject matter.

Figure A4. Text “The Homework Debate: How Homework Benefits Students”.

Appendix B Questionnaire

Informed Consent

The study for which we are asking you to participate belongs to the project Assessing digital reading skills across ages, led by Raquel Cerdán (Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020; Referencia: PID2020-117996RB-I00). The aim of the present research is to evaluate the use and effectiveness of the eye tracker to record eye movements occurring while we read.

If you agree to participate, we will ask you to fill out the questionnaire that you will find afterwards, which will take you about 5 to 10 min. Once you have answered the questions, we will contact you to schedule a date to initiate the eye tracker experiment. This session will last approximately 1 h. The study does not involve any kind of risk and your participation is completely voluntary.

The data gathered from this research will be handled confidentially and will not be used for any purpose other than for this investigation. As commented, your participation is fully voluntary, and you have the right to withdraw from the study at any time.

If you have any questions, please contact Mariola Giménez: magisal2@alumni.uv.es

I have read the previous information and I give my consent to participate in this study: Yes/No

Personal Information

1. Name and surname:
2. Email address:
3. Phone number:
4. Gender:
 - Male.
 - Female.
 - Non-binary.
 - Prefer not to say.
 - Other:
5. Age:
6. Country of origin:
 - Spain.
 - Other:

7. Highest level of education completed:
 - Upper secondary school (bachillerato).
 - Advanced vocational training (formación profesional de grado superior).
 - University degree, master's degree or PhD.
 - Other:
8. Degree access score (In the Spanish educational system, the degree access score is out of 14. If you are an Erasmus student, please indicate the degree access score you obtained in your educational system):
9. Do you have any learning difficulties related to reading and/or writing?
 - Yes.
 - No.
10. At the moment, do you have any vision problems that prevent you from reading normally? (If you use glasses or contact lenses to correct the difficulty, please answer yes only if the problem persists even when you wear the glasses or use the contact lenses)
 - Yes.
 - No.
11. Do you use bifocal, trifocal or progressive glasses/lenses? (These types of glasses correct for near and distant vision at the same time)
 - Yes.
 - No.

Homework Debate

Please answer the following questions sincerely, by indicating the degree to which you disagree or agree (from 1 to 5, with 1 indicating *minimum agreement* and 5 *maximum agreement*).

1. I am quite familiar with the current debate about homework:
2. I am quite interested in reading texts dealing with a debate about homework:
3. I believe homework is completely necessary for school learning:
4. I believe homework should be eliminated from schools (in other words, children should work during school time):
5. I think homework can help children learn better:
6. I think homework does not help the teacher in supporting the students' learning processes:

Doubts or Comments

Do you have any doubts or comments you would like to share with the research team?

Final Section

Click "Send"/"Enviar" to complete the questionnaire.

Thank you very much for your participation! We will contact you soon to schedule a date to initiate the experiment.

References

1. Stadler, M.; Scharrer, L.; Skodzik, T.; Bromme, R. Comprehending Multiple Documents on Scientific Controversies: Effects of Reading Goals and Signaling Rhetorical Relationships. *Discourse Process* **2014**, *51*, 93–116. [CrossRef]
2. Rouet, J.-F.; Britt, M.A.; Durik, A.M. RESOLV: Readers' Representation of Reading Contexts and Tasks. *Educ. Psychol.* **2017**, *52*, 200–215. [CrossRef]
3. Ministerio de Educación, Formación Profesional y Deportes. *PISA 2022. Programa para la Evaluación Internacional de los Estudiantes. Informe Español*, 1st ed.; Instituto Nacional de Evaluación Educativa: Madrid, España, 2023; pp. 68–70, EAN 9789200256226. Available online: https://www.libreria.educacion.gob.es/libro/pisa-2022-programa-para-la-evaluacion-internacional-de-los-estudiantes-informe-espanol_183950/ (accessed on 29 March 2024).
4. OECD. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*; OECD Publishing: Paris, France, 2013; pp. 199–200, ISBN 978-92-64-20425-6.
5. Stadler, M.; Bromme, R. The Content-Source Integration Model: A Taxonomic Description of How Readers Comprehend Conflicting Scientific Information. In *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and*

- the Educational Sciences*; Rapp, D.N., Braasch, J.L.G., Eds.; MIT Press: Cambridge, MA, USA, 2014; pp. 379–402, ISBN 978-0-262-32564-6.
6. Britt, M.A.; Perfetti, C.A.; Sandak, R.; Rouet, J.F. Content Integration and Source Separation in Learning from Multiple Texts. In *Narrative, Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso*; Goldman, S.R., Graesser, A.C., van den Broek, P., Eds.; Erlbaum: Mahwah, NJ, USA, 1999; pp. 209–233, ISBN 9781410603135.
 7. Maier, J.; Richter, T. Text Belief Consistency Effects in the Comprehension of Multiple Texts with Conflicting Information. *Cogn. Instr.* **2013**, *31*, 151–175. [[CrossRef](#)]
 8. Maier, J.; Richter, T. Fostering Multiple Text Comprehension: How Metacognitive Strategies and Motivation Moderate the Text-Belief Consistency Effect. *Metacogn. Learn.* **2014**, *9*, 51–74. [[CrossRef](#)]
 9. Wiley, J. A Fair and Balanced Look at the News: What Affects Memory for Controversial Arguments? *J. Mem. Lang.* **2005**, *53*, 95–109. [[CrossRef](#)]
 10. Richter, T.; Maier, J. Comprehension of Multiple Documents with Conflicting Information: A Two-Step Model of Validation. *Educ. Psychol.* **2017**, *52*, 148–166. [[CrossRef](#)]
 11. Abendroth, J.; Richter, T. Text-Belief Consistency Effect in Adolescents' Comprehension of Multiple Documents from the Web (El Efecto de Consistencia en la Comprensión Lectora de los Adolescentes de Documentos Múltiples Provenientes de Internet). *Infanc. Aprendiz.* **2020**, *43*, 60–100. [[CrossRef](#)]
 12. Mason, A.E.; Braasch, J.L.G.; Greenberg, D.; Kessler, E.D.; Allen, L.K.; McNamara, D.S. Comprehending Multiple Controversial Texts about Childhood Vaccinations: Topic Beliefs and Integration Instructions. *Read. Psychol.* **2022**, *44*, 436–462. [[CrossRef](#)]
 13. Holmqvist, K.; Nyström, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. *Eye Tracking: A Comprehensive Guide to Methods and Measures*; Oxford University Press: New York, NY, USA, 2011; ISBN 978-0-19969708-3.
 14. Rayner, K. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol. Bull.* **1998**, *124*, 372–422. [[CrossRef](#)]
 15. Jarodzka, H.; Holmqvist, K.; Gruber, H. Eye Tracking in Educational Science: Theoretical Frameworks and Research Agendas. *J. Eye Mov. Res.* **2017**, *10*, 1–8. [[CrossRef](#)]
 16. Maier, J.; Richter, T.; Britt, M.A. Cognitive Processes Underlying the Text-Belief Consistency Effect: An Eye-Movement Study. *Appl. Cogn. Psychol.* **2018**, *32*, 171–185. [[CrossRef](#)]
 17. Abendroth, J.; Richter, T. Reading Perspectives Moderate Text-Belief Consistency Effects in Eye Movements and Comprehension. *Discourse Process* **2023**, *60*, 119–140. [[CrossRef](#)]
 18. Hoover, W.A.; Gough, P.B. The Simple View of Reading. *Read. Writ.* **1990**, *2*, 127–160. [[CrossRef](#)]
 19. Britt, M.A.; Rouet, J.-F.; Durik, A.M. *Literacy beyond Text Comprehension*; Routledge: New York, NY, USA, 2017; ISBN 9781315682860.
 20. List, A.; Alexander, P.A. Toward an Integrated Framework of Multiple Text Use. *Educ. Psychol.* **2019**, *54*, 20–39. [[CrossRef](#)]
 21. OECD. *PISA 2018 Assessment and Analytical Framework*; OECD Publishing: Paris, France, 2019; pp. 21–36. ISBN 978-92-64-47759-9.
 22. Kintsch, W. *Comprehension: A Paradigm for Cognition*; Cambridge University Press: Cambridge, MA, USA, 1998; ISBN 9780521629867.
 23. McNamara, D.S.; Magliano, J. Toward a Comprehensive Model of Comprehension. In *The Psychology of Learning and Motivation*; Ross, B.H., Ed.; Elsevier Academic Press: Cambridge, MA, USA, 2009; Volume 76, pp. 297–384, ISBN 9780323990998.
 24. van den Broek, P.; Rapp, D.N.; Kendeou, P. Integrating Memory-Based and Constructionist Processes in Accounts of Reading Comprehension. *Discourse Process* **2005**, *39*, 299–316. [[CrossRef](#)]
 25. Perfetti, C.A.; Rouet, J.F.; Britt, M.A. Toward a Theory of Documents Representation. In *The Construction of Mental Representation During Reading*; Van Oostendorp, H., Goldman, S.R., Eds.; Psychology Press: New York, NY, USA, 1999; pp. 99–122, ISBN 9781410603050.
 26. Rouet, J.-F. *The Skills of Document Use*; Routledge: New York, NY, USA, 2006; ISBN 9780429235962.
 27. Nickerson, R.S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Rev. Gen. Psychol.* **1998**, *2*, 175–220. [[CrossRef](#)]
 28. Festinger, L. *A Theory of Cognitive Dissonance*; Stanford University Press: Redwood City, CA, USA, 1957; ISBN 9781503620766.
 29. Kahneman, D. Maps of Bounded Rationality: Psychology for Behavioral Economics. *Am. Econ. Rev.* **2003**, *93*, 1449–1475. [[CrossRef](#)]
 30. Sweller, J.; Ayres, P.; Kalyuga, S. *Cognitive Load Theory*; Springer New York: New York, NY, USA, 2011; ISBN 978-1-4419-8126-4.
 31. van Strien, J.L.H.; Brand-Gruwel, S.; Boshuizen, H.P.A. Dealing with Conflicting Information from Multiple Nonlinear Texts: Effects of Prior Attitudes. *Comput. Human. Behav.* **2014**, *32*, 101–111. [[CrossRef](#)]
 32. Anmarkrud, Ø.; Bråten, I.; Strømsø, H.I. Multiple-Documents Literacy: Strategic Processing, Source Awareness, and Argumentation When Reading Multiple Conflicting Documents. *Learn. Individ. Differ.* **2014**, *30*, 64–76. [[CrossRef](#)]
 33. Braasch, J.L.G.; Haverkamp, Y.E.; Latini, N.; Shaw, S.; Arshad, M.S.; Bråten, I. Belief Bias When Adolescents Read to Comprehend Multiple Conflicting Texts. *Read. Writ.* **2022**, *35*, 1759–1785. [[CrossRef](#)]
 34. Tarchi, C.; Mason, L. Effects of Critical Thinking on Multiple-Document Comprehension. *Eur. J. Psychol. Educ.* **2020**, *35*, 289–313. [[CrossRef](#)]
 35. Abendroth, J.; Richter, T. Mere Plausibility Enhances Comprehension: The Role of Plausibility in Comprehending an Unfamiliar Scientific Debate. *J. Educ. Psychol.* **2021**, *113*, 1304–1322. [[CrossRef](#)]
 36. Maier, J.; Richter, T. Effects of Text-Belief Consistency and Reading Task on the Strategic Validation of Multiple Texts. *Eur. J. Psychol. Educ.* **2016**, *31*, 479–497. [[CrossRef](#)]
 37. Patson, N.D.; Warren, T. Eye Movements When Reading Implausible Sentences: Investigating Potential Structural Influences on Semantic Integration. *Q. J. Exp. Psychol.* **2010**, *63*, 1516–1532. [[CrossRef](#)] [[PubMed](#)]

38. Matsuki, K.; Chow, T.; Hare, M.; Elman, J.L.; Scheepers, C.; McRae, K. Event-Based Plausibility Immediately Influences On-Line Language Comprehension. *J. Exp. Psychol. Learn. Mem. Cogn.* **2011**, *37*, 913–934. [\[CrossRef\]](#)
39. Staub, A.; Rayner, K.; Pollatsek, A.; Hyönä, J.; Majewski, H. The Time Course of Plausibility Effects on Eye Movements in Reading: Evidence from Noun-Noun Compounds. *J. Exp. Psychol. Learn. Mem. Cogn.* **2007**, *33*, 1162–1169. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Rinck, M.; Gámez, E.; Díaz, J.M.; De Vega, M. Processing of Temporal Information: Evidence from Eye Movements. *Mem. Cognit.* **2003**, *31*, 77–86. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Salmeron, L.; Gil, L.; Bråten, I. Using Eye-Tracking to Assess Sourcing during Multiple Document Reading: A Critical Analysis. *Frontline Learn. Res.* **2018**, *6*, 105–122. [\[CrossRef\]](#)
42. Alemdag, E.; Cagiltay, K. A Systematic Review of Eye Tracking Research on Multimedia Learning. *Comput. Educ.* **2018**, *125*, 413–428. [\[CrossRef\]](#)
43. Just, M.A.; Carpenter, P.A. A Theory of Reading: From Eye Fixations to Comprehension. *Psychol. Rev.* **1980**, *87*, 329–354. [\[CrossRef\]](#)
44. Reichle, E.D.; Pollatsek, A.; Rayner, K. E-Z Reader: A Cognitive-Control, Serial-Attention Model of Eye-Movement Behavior during Reading. *Cogn. Syst. Res.* **2006**, *7*, 4–22. [\[CrossRef\]](#)
45. Jacob, R.J.K.; Karn, K.S. Eye Tracking in Human-Computer Interaction and Usability Research. In *The Mind's Eye*; Hyönä, J., Radach, R., Deubel, H., Eds.; North Holland: Amsterdam, The Netherlands, 2003; pp. 573–605. ISBN 978-0-444-51020-4.
46. Lai, M.-L.; Tsai, M.-J.; Yang, F.-Y.; Hsu, C.-Y.; Liu, T.-C.; Lee, S.W.-Y.; Lee, M.-H.; Chiou, G.-L.; Liang, J.-C.; Tsai, C.-C. A Review of Using Eye-Tracking Technology in Exploring Learning from 2000 to 2012. *Educ. Res. Rev.* **2013**, *10*, 90–115. [\[CrossRef\]](#)
47. Jarodzka, H.; Brand-Gruwel, S. Tracking the Reading Eye: Towards a Model of Real-World Reading. *J. Comput. Assist. Learn.* **2017**, *33*, 193–201. [\[CrossRef\]](#)
48. Rayner, K. Eye Movements in Reading: Models and Data. *J. Eye Mov. Res.* **2009**, *2*, 1–10. [\[CrossRef\]](#)
49. Clifton, C.; Staub, A.; Rayner, K. Eye Movements in Reading Words and Sentences. In *Eye Movements: A Window on Mind and Brain*; van Gompel, R.P.G., Fischer, M.H., Murray, W.S., Hill, R.L., Eds.; Elsevier Science: Amsterdam, The Netherlands, 2007; pp. 341–371. ISBN 978-0-08-04498-7.
50. Winke, P.; Lim, H. ESL Essay Raters' Cognitive Processes in Applying the Jacobs et al. Rubric: An Eye-Movement Study. *Assess. Writ.* **2015**, *25*, 38–54. [\[CrossRef\]](#)
51. Salmerón, L.; Delgado, P.; Mason, L. Using Eye-Movement Modelling Examples to Improve Critical Reading of Multiple Webpages on a Conflicting Topic. *J. Comput. Assist. Learn.* **2020**, *36*, 1038–1051. [\[CrossRef\]](#)
52. Mason, L.; Zaccoletti, S.; Scrimin, S.; Tornatora, M.C.; Florit, E.; Goetz, T. Reading with the Eyes and under the Skin: Comprehending Conflicting Digital Texts. *J. Comput. Assist. Learn.* **2020**, *36*, 89–101. [\[CrossRef\]](#)
53. Gottschling, S.; Kammerer, Y. Readers' Regulation and Resolution of a Scientific Conflict Based on Differences in Source Information: An Eye-Tracking Study. *Discourse Process* **2021**, *58*, 468–490. [\[CrossRef\]](#)
54. Wang, C.-Y.; Tsai, M.-J.; Tsai, C.-C. Predicting Cognitive Structures and Information Processing Modes by Eye-Tracking When Reading Controversial Reports about Socio-Scientific Issues. *Comput. Human. Behav.* **2020**, *112*, 106471. [\[CrossRef\]](#)
55. Stadler, M.; Scharrer, L.; Bromme, R. How Relevance Affects Understanding of Conflicts Between Multiple Documents: An Eye-Tracking Study. *Read. Res. Q.* **2020**, *55*, 625–641. [\[CrossRef\]](#)
56. Voss, J.F.; Fincher-Kiefer, R.; Wiley, J.; Silfies, L.N. On the Processing of Arguments. *Argumentation* **1993**, *7*, 165–181. [\[CrossRef\]](#)
57. Saux, G.; Britt, M.A.; Vibert, N.; Rouet, J. Building Mental Models from Multiple Texts: How Readers Construct Coherence from Inconsistent Sources. *Lang. Linguist. Compass* **2021**, *15*, e12409. [\[CrossRef\]](#)
58. Cerdán, R.; Máñez, I.; Serrano-Mendizábal, M.; Richter, T.; Herrero, L. Reading Controversial Texts: Effects of Beliefs and Stakes on Undergraduates' Argument Integration. In Proceedings of the 33rd Annual Meeting of the Society for Text and Discourse, Oslo, Norway, 30 June 2023.
59. Flesch, R. A New Readability Yardstick. *J. Appl. Psychol.* **1948**, *32*, 221–233. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Tobii AB. Tobii Pro Lab User Manual (Version v 1.232). Available online: https://go.tobii.com/tobii_pro_lab_user_manual (accessed on 29 January 2024).
61. Kaakinen, J.K.; Olkonemi, H.; Kinnari, T.; Hyönä, J. Processing of Written Irony: An Eye Movement Study. *Discourse Process* **2014**, *51*, 287–311. [\[CrossRef\]](#)
62. Mason, L.; Pluchino, P.; Tornatora, M.C. Eye-Movement Modeling of Integrative Reading of an Illustrated Text: Effects on Processing and Learning. *Contemp. Educ. Psychol.* **2015**, *41*, 172–187. [\[CrossRef\]](#)
63. Eskenazi, M.A. Best Practices for Cleaning Eye Movement Data in Reading Research. *Behav. Res. Methods* **2023**, *56*, 2083–2093. [\[CrossRef\]](#)
64. Olsen, A. The Tobii I-VT Fixation Filter. Algorithm Description. Available online: https://stemedhub.org/resources/2173/download/Tobii_WhitePaper_TobiiIVTFixationFilter.pdf (accessed on 29 January 2024).
65. Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 5th ed.; SAGE: London, UK, 2017; ISBN 978-1-5264-1951-4.
66. OECD. *PISA 2022 Results (Volume I)*; OECD Publishing: Paris, France, 2023; ISBN 978-92-64-35128-8.
67. Ayroles, J.; Potocki, A.; Ros, C.; Cerdán, R.; Britt, M.A.; Rouet, J. Do You Know What You Are Reading for? Exploring the Effects of a Task Model Enhancement on Fifth Graders' Purposeful Reading. *J. Res. Read.* **2021**, *44*, 837–858. [\[CrossRef\]](#)
68. Stark, L.; Korbach, A.; Brünken, R.; Park, B. Measuring (Meta)Cognitive Processes in Multimedia Learning: Matching Eye Tracking Metrics and Think-Aloud Protocols in Case of Seductive Details. *J. Comput. Assist. Learn.* **2024**, 1–20. [\[CrossRef\]](#)

69. Carr, J.W.; Pescuma, V.N.; Furlan, M.; Ktori, M.; Crepaldi, D. Algorithms for the Automated Correction of Vertical Drift in Eye-Tracking Data. *Behav. Res. Methods* **2022**, *54*, 287–310. [\[CrossRef\]](#)
70. Karimi, M.N.; Ashkani, P. Situation-Model Representations of Conflicting Textual Information in L2 Readers: The Effects of Prior Beliefs and L2 Proficiency. *Lang. Aware.* **2023**, *32*, 323–341. [\[CrossRef\]](#)
71. Toulmin, S.E. *The Uses of Argument*; Cambridge University Press: Cambridge, UK, 2003; ISBN 9780511840005.
72. Talwar, A.; Magliano, J.P.; Higgs, K.; Santuzzi, A.; Tonks, S.; O'Reilly, T.; Sabatini, J. Early Academic Success in College: Examining the Contributions of Reading Literacy Skills, Metacognitive Reading Strategies, and Reading Motivation. *J. Coll. Read. Learn.* **2023**, *53*, 58–87. [\[CrossRef\]](#)
73. Magliano, J.P.; Talwar, A.; Feller, D.P.; Wang, Z.; O'Reilly, T.; Sabatini, J. Exploring Thresholds in the Foundational Skills for Reading and Comprehension Outcomes in the Context of Postsecondary Readers. *J. Learn. Disabil.* **2023**, *56*, 43–57. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Kalder, G.; Higgs, K.; Lampi, J.; Santuzzi, A.; Tonks, S.M.; O'Reilly, T.; Sabatini, J.P.; Magliano, J.P. Testing the Model of a Proficient Academic Reader (PAR) in a Postsecondary Context. *Read. Writ.* **2024**, 1–40. [\[CrossRef\]](#)
75. Cromley, J.G.; Azevedo, R. Self-Report of Reading Comprehension Strategies: What Are We Measuring? *Metacogn Learn.* **2007**, *1*, 229–247. [\[CrossRef\]](#)
76. Tobii AB. Tobii Pro Glasses 3 User Manual. Available online: <https://go.tobii.com/tobii-pro-glasses-3-user-manual> (accessed on 6 June 2024).
77. Wiggins, G. *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*; Jossey-Bass Publishers: San Francisco, CA, USA, 1998; ISBN 0-7879-0848-7.
78. McNamara, D.S. SERT: Self-Explanation Reading Training. *Discourse Process* **2004**, *38*, 1–30. [\[CrossRef\]](#)
79. McNamara, D.S.; Scott, J.L. Training Self Explanation and Reading Strategies. In Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, Houston, TX, USA, 27 September–1 October 1999; Volume 43, pp. 1156–1160.
80. O'Reilly, T.; Best, R.; McNamara, D.S. Self-Explanation Reading Training: Effects for Low-Knowledge Readers. In Proceedings of the 26th Annual Meeting of the Cognitive Science Society, Chicago, IL, USA, 4–7 August 2004; Forbus, K., Gentner, D., Reiger, T., Eds.; Erlbaum: Mahwah, NJ, USA, 2004; Volume 26, pp. 1053–1058.
81. McNamara, D.S.; Levinstein, I.B.; Boonthum, C. iSTART: Interactive Strategy Training for Active Reading and Thinking. *Behav. Res. Methods Instrum. Comput.* **2004**, *36*, 222–233. [\[CrossRef\]](#)
82. McNamara, D.S.; O'Reilly, T.P.; Best, R.M.; Ozuru, Y. Improving Adolescent Students' Reading Comprehension with iSTART. *J. Educ. Comput. Res.* **2006**, *34*, 147–171. [\[CrossRef\]](#)
83. Magliano, J.P.; Todaro, S.; Millis, K.; Wiemer-Hastings, K.; Kim, H.J.; McNamara, D.S. Changes in Reading Strategies as a Function of Reading Training: A Comparison of Live and Computerized Training. *J. Educ. Comput. Res.* **2005**, *32*, 185–208. [\[CrossRef\]](#)
84. Bandura, A. *Social Learning Theory*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1977; ISBN 0-13-816751-6.
85. van Marlen, T.; van Wermeskerken, M.; Jarodzka, H.; van Gog, T. Effectiveness of Eye Movement Modeling Examples in Problem Solving: The Role of Verbal Ambiguity and Prior Knowledge. *Learn. Instr.* **2018**, *58*, 274–283. [\[CrossRef\]](#)
86. Mason, L.; Pluchino, P.; Tornatora, M.C. Using Eye-Tracking Technology as an Indirect Instruction Tool to Improve Text and Picture Processing and Learning. *Br. J. Educ. Technol.* **2016**, *47*, 1083–1095. [\[CrossRef\]](#)
87. Salmerón, L.; Llorens, A. Instruction of Digital Reading Strategies Based on Eye-Movements Modeling Examples. *J. Educ. Comput. Res.* **2019**, *57*, 343–359. [\[CrossRef\]](#)
88. Mason, L.; Scheiter, K.; Tornatora, M.C. Using Eye Movements to Model the Sequence of Text–Picture Processing for Multimedia Comprehension. *J. Comput. Assist. Learn.* **2017**, *33*, 443–460. [\[CrossRef\]](#)
89. Li, X. Eye-Tracking Research in Interactive Language Learning Environments: A Systematic Review. *Educ. Inf. Technol.* **2024**, 1–26. [\[CrossRef\]](#)

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.