What you read vs. what you know: a methodologically diverse approach to unraveling the neurocognitive architecture of text-based and knowledge-based validation processes during reading

Moort, M.L. van

Citation

Moort, M. L. van. (2022, March 3). *What you read vs. what you know: a methodologically diverse approach to unraveling the neurocognitive architecture of text-based and knowledge-based validation processes during reading*. Retrieved from https://hdl.handle.net/1887/3278025

5

Purposeful validation: Are validation processes and the construction of a mental representation influenced by reading goal?

This chapter is based on:
Van Moort, M.L., Koornneef A., & van den Broek, P.W. (under revision). Purposeful validation: Do reading goals affect monitoring processes during reading and the construction of a mental representation? *Journal of Educational Science.*

Abstract

People read for many different reasons. These goals affect the cognitive processes and strategies they use during reading. Understanding how reading goals exert their effects requires investigation of whether and how they affect specific component processes, such as validation. We investigated the effects of reading goal on text-based and knowledge-based validation processes during reading and on the resulting offline mental representation. We employed a self-paced sentence-by-sentence contradiction paradigm with versions of texts containing target sentences that varied systematically in congruency with prior text and accuracy with background knowledge. Participants were instructed to read for general comprehension or study. Memory for text information was assessed the next day. We also measured the degree to which each text topic was novel to a reader and their working-memory capacity. Results show that reading goals affect readers' general processing but provide no clear evidence that reading goals influence online validation processes. However, reading goal effects on readers' memory for target information did depend on the congruency of that information with the preceding text: Reading for study generally resulted in better memory for target information than reading for comprehension did, but not for target information that was incongruent with prior text. These results suggest that reading goals may not influence validation processes directly but affect the processes that take place after the detection of an inconsistency – particularly in the case of incongruencies with prior text.

Keywords: comprehension, reading goals, monitoring, background knowledge, memory, validation

Introduction

Reading is a purposeful activity. People can have many different reasons for reading a text: they can read for pleasure, to learn for school, to obtain instructions, and so on. It is clear that these different goals affect the cognitive processes and strategies readers use when they proceed through a text (Britt et al., 2018; Linderholm et al., 2004; McCrudden et al., 2011; van den Broek et al., 1999, 2001). Changes in cognitive processes, in turn, affect readers' memory for text information (Lorch et al., 1993, 1995; van den Broek et al., 2001), particularly memory for text information that is relevant to their reading purpose (Anderson, Pichert, & Shirey, 1983; Baillet & Keenan, 1986; Hyönä, Lorch, Kaakinen, Lorch Jr, & Kaakinen, 2002; Pichert & Anderson, 1977).

To understand how reading goals modulate reading process and outcomes, it is necessary to investigate how they affect specific component processes that occur during reading. Successful comprehension requires readers to continually use various forms of information – for example, semantic (the meaning of words), syntactic (grammatical), and pragmatic (their understanding of the world) – to build a coherent, meaningful mental representation or situation model of a text (Graesser et al., 1994; Johnson-Laird, 1983; Kintsch, 1988; van den Broek, 1988; Zwaan & Singer, 2003). An essential aspect of building such representations is that readers monitor to what extent the incoming information is both coherent with prior information in the text (i.e., congruent) and valid with respect to their background knowledge (i.e., accurate) – a process called validation (O'Brien & Cook, 2016a; Richter & Rapp, 2014; Singer, 2013, 2019; Singer et al., 1992). By validating incoming information readers establish coherence during reading and protect the emerging mental representation against incongruencies and inaccuracies.

In the current study we investigated whether a reader's purpose for reading affects validation processes and, if so, in what manner. As described below, prior studies on the influence of reading purpose on comprehension involved examinations of how people read and learn valid, accurate information (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; Narvaez et al., 1999; Salmerón et al., 2010; van den Broek et al., 2001; Yeari et al., 2015). But in daily life people are not always presented with accurate information; they frequently encounter ideas and concepts that are inaccurate or incongruent, representing misinformation or even fake news (Richter & Rapp, 2014). Therefore, it is crucial that we understand how reading purpose may affect the processing of texts containing false or incongruent information and readers' subsequent memory for those texts. To this end, we considered the influence of reading goals on validation processes during reading as well as their effects on the final product of reading a text (i.e., the offline memory representation).

Purposeful reading

There is considerable evidence that readers' purpose for reading affects general reading processes and comprehension (Britt et al., 2018; Cain, 1999; Kaakinen & Hyönä, 2005, 2010; Narvaez et al., 1999; van den Broek et al., 2001; Van den Broek et al., 1995). The development of such reading goals is generally assumed to be influenced by the instructions provided by the reader – either directly or in interaction with the personal intentions of the reader (e.g., McCrudden, Magliano, & Schraw, 2010; McCrudden & Schraw, 2007). Such instructions can highlight discrete text elements by posing pre-reading questions or objectives (e.g., by prompting readers to identify specific text segments), prompt individuals to read a text from a designated reference point (e.g., from a particular perspective) or prompt individuals to read for a general purpose (e.g., reading for study, reading for entertainment, reading for general comprehension) (Bohn-Gettler & Kendeou, 2014; Bråten & Samuelstuen, 2004; Cerdán & Vidal-Abarca, 2008; Linderholm & van den Broek, 2002; Narvaez et al., 1999; Rouet et al., 2001; Salmerón et al., 2010; van den Broek et al., 2001; Yeari et al., 2015). For example, in some studies readers were asked to read a text describing a location and evaluate whether that location would be suitable for living (e.g., Hyönä et al., 2002; McCrudden et al., 2010; McCrudden & Schraw, 2007) or to read a text describing a house from the perspective of a potential home buyer or a burglar (e.g., Baillet & Keenan, 1986; McCrudden, Schraw, & Kambe, 2005). In other studies, more general instructions were given to modify readers' criteria for how well or how deeply they should process the text, for example by contrasting superficial or lower-effort reading purposes (e.g., reading for pleasure or proofreading) with deeper or more effortful reading purposes (e.g., reading in preparation for an exam) (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; Narvaez et al., 1999; Salmerón et al., 2010; van den Broek et al., 2001).

Results of these studies show that the way people read and what information they acquire systematically varies as a function of reading purpose. These studies investigated various aspects of reading comprehension. Some studies have focused on the effects of reading goals on the cognitive processes that take place during reading (i.e., online), whereas others have focused on the outcome of comprehension (i.e., the offline representation). Online studies have shown that relevance instructions affect readers' attention toward relevant and irrelevant information (Goetz et al., 1983; Kaakinen et al., 2002; McCrudden et al., 2005). Furthermore, readers with more effortful general reading purposes (e.g., reading for study) spend more time reading the texts (Yeari et al., 2015) and engage in more coherence-building processes during reading (e.g., generating connecting, explanatory, and predictive inferences) than readers with superficial or lower-effort reading purposes (e.g., reading for etertainment) (Linderholm & van den Broek, 2002; Lorch et al., 1993; Narvaez et al., 1999; van den Broek et al., 2001). In general, it seems that demanding reading purposes lead to more careful text processing than superficial reading purposes (Lorch et al., 1993, 1995; van den Broek et al., 2001). With respect to the offline mental representation, more demanding reading purposes result in the construction of a more coherent text representation and in better memory for the text than superficial reading purposes (Britt et al., 2018; Linderholm et al., 2004; Lorch et al., 1993, 1995; van den Broek et al., 2001; Yeari et al., 2015).

Validating mental representations

To protect the mental representation of a text against incongruencies and inaccuracies, readers routinely validate incoming information against various sources of information – most notably the preceding text and their own background knowledge (Isberner & Richter, 2014a; Nieuwland & Kuperberg, 2008; O’Brien & Cook, 2016a, 2016b; Schroeder et al., 2008; Singer, 2006). In describing the cognitive architecture of validation, theoretical models assume distinct components of validation: a coherence-detection component and a post-detection processing component (Cook & O’Brien, 2014; Isberner & Richter, 2014a; Richter, 2015; Singer, 2019; van den Broek & Helder, 2017). The coherence-detection component, involved in detecting (in)consistencies, is the main focus of the RI-Val model of comprehension (Cook & O’Brien, 2014; O’Brien & Cook, 2016a, 2016b). In this model, validation is described as one of three processing stages – resonance, integration, and validation – that comprise comprehension. According to the model, incoming information activates related information from long-term memory via a low-level passive resonance mechanism (Myers & O’Brien, 1998; O’Brien & Myers, 1999). This activated information then is integrated with the contents of working memory and these linkages made during the integration stage are validated against information in memory that is readily available to the reader (i.e., information that either already is part of working memory or easily can be made available from long-term memory) in a single, passive pattern-matching process (e.g., McKoon & Ratcliff, 1995; Myers & O’Brien, 1998; O’Brien & Albrecht, 1992). These contents of active memory include both portions of the episodic representation of the text (i.e., context) and general world knowledge. In addition, the model includes a coherence threshold: a point at which processing is deemed ‘good enough’ for the reader to move on in a text. This threshold is assumed to be flexible: readers may wait for more or less information to accrue before moving on in the text depending on variables associated with the reader, the task and the text (O’Brien & Cook, 2016b). The three processes are assumed to have an asynchronous onset and to run to completion over time, regardless of whether the reader has moved on in the text (i.e., reached their coherence threshold).

Once detected, inconsistencies may trigger further processing. Such post-detection processes include possible efforts to repair coherence triggered by the detection of the inconsistency as elaborated in a second validation model, the two-step model of validation (Isberner & Richter, 2014a; Richter, 2011; Richter et al., 2009; Schroeder et al., 2008). In this model, validation is described as consisting of two components: (1) epistemic monitoring (i.e., detecting inconsistencies) during a comprehension stage, followed by (2) optional epistemic elaboration processes (e.g., resolving inconsistencies) during an evaluative stage (e.g., Isberner & Richter, 2014; Richter, 2011; Richter, Schroeder, & Wöhrmann, 2009; Schroeder et al., 2008). According to this model the initial detection of inconsistencies (i.e., epistemic monitoring) is a routine part of comprehension. Similar to the RI–Val model, these detection processes are memory-based, pose little demands on cognitive resources, and are not dependent on readers' goals (Richter et al., 2009). If an inconsistency is detected, readers may initiate epistemic elaboration processes to resolve the inconsistency. Such elaboration processes may take place during reading (e.g., generating elaborative and bridging inferences to establish hypothetical truth conditions) or after reading of a text is completed (e.g., searching for evidence that could support dubious information). These processes only occur when readers are motivated and have enough cognitive resources available, as these processes are assumed to be slow, resource-demanding and under strategic control of the reader (Maier & Richter, 2013; Richter, 2011).

Theoretical accounts such as the two-step model of validation emphasize that validation processes function as a gatekeeper for the quality of the mental representation of a text and as such assume a close relation between online validation processes and offline memory products (e.g., Isberner & Richter, 2014; Singer, 2006, 2019). On the one hand, successful validation (i.e., information is deemed congruent and accurate) should result in the integration of the incoming information into the emerging mental representation thereby increasing the likelihood that it will be encoded in readers' long-term memory (Schroeder et al., 2008; Singer, 2006, 2019). On the other hand, if validation fails (i.e., incoming information is deemed inaccurate or incongruent), integration of the incoming information into the reader's mental representation and long-term memory fails – making this information harder to remember. Consistent with this idea, readers tend to have poorer memory for inaccurate or incongruent text information than for accurate or congruent text information (e.g., Schroeder et al., 2008; van Moort, Jolles, et al., 2020).

As described, current theoretical frameworks offer a time course and a rudimentary cognitive architecture for validation processes. Furthermore, they generally agree that incoming information is routinely validated against a reader's evolving situation model of a text (e.g., Isberner & Richter, 2014; Nieuwland & Kuperberg, 2008; O’Brien & Cook, 2016a, 2016b; Schroeder et al., 2008; Singer et al., 1992; Singer, 2006, 2013). Because a situation model comprises both textual and world knowledge information, most accounts assume that both sources can affect validation processes, yet few accounts make an explicit distinction between these sources in their depiction of the cognitive architecture of validation. Recent research shows that such distinction is essential as incoming information may be validated against contextual information and background knowledge through dissociable, interactive, validation channels involving (partially) distinct neurocognitive mechanisms (van Moort et al., 2018, 2020, 2021). Furthermore, although readers are assumed to use both sources of information for validation, the dominance of one informational source over the other may depend on the strength of the reader’s topic-relevant world knowledge (Cook & O’Brien, 2014) versus the strength of the contextual information (Cook & Guéraud, 2005; Myers, Cook, Kambe, Mason, & O’Brien, 2000; O’Brien & Albrecht, 1991).

Validation and reading goals

The degree to which readers validate incoming information may depend on their purpose for reading (Singer, 2019). Studies have shown that readers' sensitivity to false or implausible information varies with their goals (e.g., Rapp, Hinze, Kohlhepp, & Ryskin, 2014). For example, Rapp et al. (2014) presented participants with stories containing both accurate and inaccurate assertions while manipulating the instructions. Participants were asked to read for comprehension or to engage in evaluative activities (e.g., fact checking and immediately correcting erroneous content or highlighting inaccuracies without changing the content). Instructions that promote evaluative activities reduced the intrusive effects of misinformation on post-reading tasks (e.g., judging the validity of statements), as compared to the performance of participants who merely read the text for comprehension.

The potential role of reading goals in online validation processes may take several forms, depending on when reading goals assert their influence. First, they may affect coherence-detection processes. In general, the RI-Val model (Cook & O’Brien, 2014; O’Brien & Cook, 2016a, 2016b) and the two-step model of validation (Isberner & Richter, 2014; Richter, 2015; Schroeder et al., 2008) do not predict strong effects of reading goals on the coherence-detection component of validation (i.e., epistemic monitoring) because both accounts assume that coherence detection involves routine processes that are, by and large, not under strategic control of the reader. However, in terms of the RI-Val model people that read for study may set a higher coherence threshold (i.e., set it later in time) accumulating more ‘evidence’ from the validation process before deeming information (in)consistent and continuing to the next sentence. If so, reading will be slower for people with a study goal and there is a greater chance that validation is complete at the end of reading a sentence. As a result, the chance that validation processes continue while reading subsequent sentences of a text (i.e., spill over) decreases. Second, reading goals may affect post-detection epistemic elaboration processes (cf. Isberner & Richter, 2014; Richter, 2015; Schroeder et al., 2008). Reading for study may result in investing more effort and resources in resolving inconsistencies than does reading for general comprehension. These elaborative repair processes can be performed immediately after the detection of an inconsistency – thereby inflating the processing times of incoherent sections of a text – but they may also (or still) be carried out after a text has been read. Consequently, the influence of reading goals may manifest itself early or late in the epistemic elaboration phase. Third, it is possible that reading goals do not influence the manner in which readers validate incoming information – neither in the coherence-detection phase nor in the epistemic elaboration phase – and only influence post-validation processes such as consolidating the newly read information in memory.

In reflecting on these options, it should be emphasized that all possible effects of reading goals on validation may vary depending on the source of a violation. For example, reading for study may focus readers more on the text itself or, alternatively, encourage them to recruit more relevant background knowledge into the mental representation. Hence, knowledge inaccuracies or contextual incongruencies (or both) may become more salient due to a study reading goal, resulting in strengthening of any observed effects.

As mentioned above, reading goals may also affect post-validation or offline memory products. Prior research shows that reading for study generally results in better memory for congruent and accurate text information (e.g., Lorch et al., 1995, 1993; van den Broek et al., 2001; Yeari et al., 2015) but whether the same holds for incongruent and inaccurate text information is unknown. Given that the quality of the offline text representation is assumed to be influenced by the cognitive processes that readers perform online (Goldman & Varma, 1995; Kintsch, 1988; Trabasso & Suh, 1993; van den Broek et al., 1999; Zwaan & Singer, 2003), comparing online and offline patterns may provide insight into the underlying mechanisms. For example, if reading for study triggers more extensive attempts to integrate the inconsistency into the representation, you would expect readers to show increased online processing difficulty and better memory for the inconsistent information. However, if reading for study triggers more thorough validation processes the detection and correction of the inconsistency may hinder the integration of the inconsistent information into the representation. If so, you would expect increased online processing difficulty and poorer memory for the inconsistent information.

Reader factors and validation

The degree to which the information in a text is novel to the reader plays a critical role in many online comprehension processes, including inference making (Cain et al., 2001; Singer, 2013), comprehension monitoring (Richter, 2015), and validation processes (e.g., Singer, 2019). With respect to validation, the more novel the information is to the reader, the less knowledge the reader has against which to validate the accuracy of the textual information. Thus, novelty likely affects knowledge-based validation processes in particular: Disruptions due to inaccurate information will be stronger when a reader is highly familiar with a topic than when most of the information in a text is new to the reader. In contrast, topic novelty is likely to either have little impact on the degree of conflict readers experience when they encounter contextual violations (i.e., incongruencies), or to have a reverse impact: Readers that lack sufficient topic-relevant knowledge may validate primarily against contextual information, resulting in stronger disruptions for textual incongruencies.

A reader's memory for a text may also be affected by the degree to which the reader has topic knowledge. In general, knowledge about a topic facilitates encoding of new information into a long-term memory representation, resulting in better memory for text information (e.g., Alexander, Kulikowich, & Schulze, 1994; Royer, Carlo, Dufresne, & Mestre, 1996; Schneider, Körkel, & Weinert, 1989; Voss & Bisanz, 1985). This knowledge effect may influence memory for all text information, irrespective of its accuracy or congruency, or only memory for accurate or congruent information. For inaccurate or incongruent information, having prior knowledge about a topic may have no positive effect or it may even impede memory, as conflicting information in the existing knowledge base may interfere with the encoding and/or retrieval of the inconsistent information. These scenarios may apply to both text and knowledge violations but perhaps most to knowledge violations, given that processing of text violations depends less on readers' prior knowledge.

In addition to topic novelty, individual differences in working-memory capacity may affect validation processes (Singer & Doering, 2014). Working memory constrains the cognitive resources available to the reader for information processing and storage (Baddeley, 1998; Baddeley & Hitch, 1974; Cowan, 1988, 2017). In the context of our study, it limits the amount of information that is available for the validation process (Hannon & Daneman, 2001; Singer, 2006) and, thus, may interfere with the ability to detect and resolve inconsistencies while reading a text. As a result, it may create a bottleneck during validation processes that may manifest itself in different ways. On the one hand, if the bottleneck primarily affects the detection of inconsistencies (i.e., coherence-detection phase), readers with a lower working-memory capacity may experience less disruption due to inconsistent information in a text than readers with a higher working-memory capacity, because lower working-memory capacity readers are less likely to detect the inconsistency. On the other hand, if the bottleneck primarily affects the repair processes that are triggered by inconsistencies (i.e., epistemic elaboration phase), readers with a lower working-memory capacity may experience more disruption due to inconsistent information than readers with a higher working-memory capacity, because lower working-memory capacity readers may have relatively fewer resources to execute the necessary inconsistency resolution. Finally, the impact of reading goals on comprehension processes in general depends on readers' working-memory capacity (Linderholm & van den Broek, 2002; Narvaez et al., 1999; van den Broek et al., 1993, 2001), and this may apply to validation processes as well.

The current study

The current study aims to provide insight into how reading goals affect validation processes and products. We compare potential online and offline effects of reading for study (readers were instructed to memorize the text information as their memory for the text contents would be tested – a commonly used high-effort reading purpose) with reading for general comprehension (readers were instructed to read for general comprehension and unaware of the memory test – the most commonly investigated purpose and the default assumption for reading comprehension models) (Kendeou et al., 2011). In addition, we examine how online validation processes are translated into offline memory representations. Because text-based and knowledge-based validation processes may be partially distinct (van Moort et al., 2018, 2020, 2021), we distinguish between these sources in our examinations. Participants read expository texts that either did or did not contain information that conflicted with the preceding text and/or readers' background knowledge (based on Van Moort et al., 2018) in a self-paced sentence-by-sentence reading task. Reading times were recorded as a measure of readers' difficulty integrating statements into a mental representation as texts unfold (Albrecht & O'Brien, 1993; Cook et al., 1998b). To assess post-reading memory, we employed a recognition memory task the following day. A secondary aim was to investigate whether potential effects of reading goals on validation processes and products were modulated by the degree to which the text topic is novel to a reader and by individual differences in working-memory capacity. As participants' knowledge on the information presented in a text may vary across texts, we asked participants to indicate for each text (immediately after reading the text) how much of the information in that text was novel to them. To investigate the possible effect of individual differences in working-memory capacity we used the

Swanson Sentence Span task (Swanson et al., 1989) as a measure of participants' working-memory capacity.

Methods

Participants

One hundred and twenty undergraduate students that were native speakers of Dutch (25 men, 95 women) aged 18-34 years (*M* = 21.6, *SD* = 3.13) participated for monetary compensation. All participants had normal or corrected-to-normal eyesight and none had diagnosed reading or learning disabilities. Participants provided written informed consent prior to testing and received financial compensation for participating. All procedures were approved by the Leiden University Institute of Education and Child Studies ethics committee and conducted in accordance with the Declaration of Helsinki.

Materials

We used the texts of van Moort, Jolles, et al. (2020) (Rapp, 2008). The forty texts are about well-known historical topics. All texts were on different topics and the contents of the texts was not related. The texts were normed to ensure that the presented facts were common knowledge in our sample (see Chapter 2 for a more detailed description of the norming study). Each text contained a target that is either true or false with respect to the readers' background knowledge; at the same time the target could either be supported or called into question by the preceding text. Hence, the context could bias towards either the true or the false target, making it either congruent or incongruent with the target (see sample text in Table 5.1). Four different versions of each of the 40 texts were constructed, by orthogonally varying the accuracy of the target with background knowledge (i.e., true/false) and the congruency of the target with the preceding context (i.e., congruent/incongruent). It is important to note that contexts biasing towards false targets did not include erroneous information; although the phrasing of the context sentences called into question the certainty of events stated in the target, all facts described in the context sentences were historically correct.

Table 5.1. Sample text with the four text versions (translated from Dutch original)

		Knowledge accuracy	
		Target true	**Target false**
Text congruency	**Target congruent with context**	**[Introduction]** In 1865, a Frenchman named Laboulaye wished to honor democratic progress in the U.S. He conceptualized a giant sculpture along with artist Auguste Bartholdi.	**[Introduction]** In 1865, a Frenchman named Laboulaye wished to honor democratic progress in the U.S. He conceptualized a giant sculpture along with artist Auguste Bartholdi.
		[Bias True Context] Their 'Statue of Liberty' would require extensive fundraising work. They organized a public lottery to generate support for the sculpture. American businessmen also contributed money to build the statue's base. Despite falling behind schedule, the statue was completed. The statue's base was finished as well and ready for mounting.	**[Bias False Context]** Their 'Statue of Liberty' would require extensive fundraising work. Raising the exorbitant funds for the statue proved an enormous challenge. Because of financial difficulties France could not afford to make a gift of the statue. Fundraising was arduous and plans quickly fell behind schedule. Because of these problems, completion of the statue seemed doomed to failure.
		[Target True] The Statue of Liberty was delivered from France to the United States.	**[Target False]** The Statue of Liberty was not delivered from France to the United States.
		[Coda] The intended site of the statue was a port in New York harbor. This location functioned as the first stop for many immigrants coming to the U.S.	**[Coda]** The intended site of the statue was a port in New York harbor. This location functioned as the first stop for many immigrants coming to the U.S.
	Target incongruent with context	**[Introduction]** In 1865, a Frenchman named Laboulaye wished to honor democratic progress in the U.S. He conceptualized a giant sculpture along with artist Auguste Bartholdi.	**[Introduction]** In 1865, a Frenchman named Laboulaye wished to honor democratic progress in the U.S. He conceptualized a giant sculpture along with artist Auguste Bartholdi.
		[Bias False Context] Their 'Statue of Liberty' would require extensive fundraising work. Raising the exorbitant funds for the statue proved an enormous challenge. Because of financial difficulties France could not afford to make a gift of the statue. Fundraising was arduous and plans quickly fell behind schedule. Because of these problems, completion of the statue seemed doomed to failure.	**[Bias True Context]** Their 'Statue of Liberty' would require extensive fundraising work. They organized a public lottery to generate support for the sculpture. American businessmen also contributed money to build the statue's base. Despite falling behind schedule, the statue was completed. The statue's base was finished as well and ready for mounting.
		[Target True] The Statue of Liberty was delivered from France to the United States.	**[Target False]** The Statue of Liberty was not delivered from France to the United States.
		[Coda] The intended site of the statue was a port in New York harbor. This location functioned as the first stop for many immigrants coming to the U.S.	**[Coda]** The intended site of the statue was a port in New York harbor. This location functioned as the first stop for many immigrants coming to the U.S.

Each text consisted of ten sentences (see Table 5.1). Sentences 1-2 were identical across conditions and introduced the topic. Sentences 3-7 differed in content, depending on context condition (congruent/incongruent). On average, the bias-true context consisted of 64 words (*SD* = 4) and 400 characters (*SD* = 22) and the bias-false context consisted of 66 words (*SD* = 4) and 406 characters (*SD* = 27). Sentence 8 was the target sentence, which was either true or false. Overall, targets were equated for length: true and false targets contained on average 9 words (*SD* = 2) and 60 characters (SD_{true} = 11; SD_{false} = 10). Half of the true targets and half of the false targets included the word not/never and half did not. Accuracy of the targets would be manipulated by either adding or omitting negation. Sentences 9-10 were identical across conditions. Sentence 9 was the spill-over sentence and did not elaborate on the fact potentially called into question in the target. Sentence 10 concluded the text. On average, texts contained 121 words (*SD* = 5) and 763 characters (*SD* = 37), across all four text versions.

To implement a repeated-measures design we used a Latin square to construct four lists, with each text appearing in a different version as a function of text context (congruent or incongruent with target) and target (true or false) on each list. The order of the texts was randomized. Each participant received one list and, hence, read one version of each text.

Experimental tasks

Reading task

Participants read the 40 texts in two blocks. Texts were presented sentence-by-sentence, while reading times were recorded. The presentation rate was self-paced and sentences remained on screen for a maximum of 10s. A fixation cross (1000 ms) was presented between texts.

At the start of the reading task, participants were instructed to read for study ("Read the texts attentively. It is important that you memorize the information in the texts, as your memory for their contents will be tested tomorrow") or to read for general comprehension ("Read the texts attentively"). Participants that were instructed to read for general comprehension were unaware of the memory test and were told that they had to perform additional cognitive tests during the second session that were part of another experiment. Participants were reminded of the instructions between blocks.

Novelty rating

After reading each text, participants indicated how much of the information in the text they just read was new to them on a visual analog scale: This scale was presented as a horizontal 100 mm line on which the novelty of the information in the text is represented by a point between the extremes of ‘nothing is new’ and ‘everything is new’. Participants’ response on this scale provides a score ranging from 0 (*nothing is new*) to 100 (*everything is new*) and provides an indication of how familiar they were with the contents of each text.

Recognition memory task

The recognition memory task (based on van Moort et al., 2020) consisted of 160 items (40 target, 40 context, 40 neutral, and 40 distractor items) that were presented in random order. Participants were presented with single sentences containing information that either matched or mismatched the information they encountered in the reading task (e.g., when they were presented with the information that the statue of liberty was delivered to the US during the reading task they could be presented with information stating either that the statue of liberty was delivered to the US or that it was not delivered to the US). The sentences that were presented in the memory task were not the exact sentences that were presented during the reading task. They were adapted to make them comprehensible outside the context of the text (e.g., anaphoric references were replaced with the original antecedent to facilitate sentence comprehension). Participants were instructed to base their answers on the information presented in the sentences, not on whether they had seen this exact sentence before. For each sentence participants indicated whether they recognized the information from the texts they read the day before (yes/no). Half of the recognition items were consistent with the version that was presented in the reading task (correct response ‘yes’), the other half was not (correct response ‘no’). Half of the presented items contained the word ‘not’ or ‘never’ and half did not (both for true and false items). Half of the recognition items were from context versions that were presented in the reading task; the other half were from the other context version. Thus, correct recognition responses included correct hits (sentence was present during the reading task and participants indicated that they read the sentence) and correct rejections (sentence was not present during the reading task and participants indicated that they did not read the sentence). Neutral sentences were presented in the reading task and stemmed from neutral parts of the text (i.e., sentence 1, 2, 9 or 10). Distractor sentences were sentences that had not been presented in the reading task.

Measures

Working-memory capacity

Working-memory capacity was measured with a Dutch version of the Swanson Sentence Span task (Swanson et al., 1989). In this task, the experimenter reads out sets of sentences, with set length increasing from 1 to 6 sentences as the test progresses. At the end of each set a comprehension question is asked about one of the sentences in the set. Participants have to remember the last word of each sentence and recall these after answering the comprehension question. The test is terminated when participants incorrectly recall a set of words or give an incorrect answer to the comprehension question twice in one set. Participants earn 0.25 points for each correct answer on the comprehension questions and each correctly recalled set of words. The sum of these points (ranging between 0-5) is the index of working-memory capacity.

Procedure

Participants were tested individually in two sessions. In the first session they completed the reading task (max. 60 minutes). After reading each text they provided a novelty rating for that text. In the second session, that took place about 24 hours after the first, they completed the recognition memory task (10-15 minutes), followed by the Swanson Sentence Span Task (max. 5 minutes) and various additional cognitive tests that were not part of the current experiment.

Analyses

To investigate the effects of the manipulations on the reading process we conducted mixed-effects linear regression analyses on the log-transformed reading times on target and spill-over sentences (i.e., sentences 8 and 9) and mixed-effects linear logistic regression on memory performance scores (i.e., probability correct) for targets using the R package LME4 version 1.1.21 (Bates et al., 2015).

For each measure we tested a model that included the random factors subjects and items and the following fixed factors: the main effects of our experimental manipulations goal (study / comprehension), accuracy (target true / false), congruency (target congruent / incongruent with context) and their interactions. In addition, we included the main effect of novelty (the amount of novel information per text, individual scores were median-centered) and the interactions between our experimental manipulations and novelty. Finally, we included the main effect of working memory capacity (individual scores were median-centered) and the interactions between our experimental manipulations and working memory capacity. Sum coding was applied in the main analyses (comprehension was coded as -0.5 and study was coded as 0.5; true was coded as -0.5 and false as 0.5; congruent was coded as -0.5 and incongruent as 0.5). For each model, residuals were normally distributed, and variance of the random effects residuals was equal across groups for subjects and items. We report the relevant fixed-effects estimates and the associated t-values (for the continuous dependent variables) and z-values (for the categorical dependent variables) in tables (see Table 5.4, 5.5, and 5.6). For ease of interpretation, we report raw means and standard errors (in ms) for relevant main effects (in text) and back-transformed estimates for interactions on a secondary y-axis (in figures). To obtain fixed-effect estimates and the associated statistics for the relevant simple effects of an interaction, pairwise comparisons were performed using the EMMEANS package (version 1.4.4) in R. In these comparisons continuous variables were centered on scores one standard deviation above and below the mean, respectively. We report odds ratio's (OR) as indices of effect size for logistic mixed models and estimated effect sizes (Cohen's d) for differences in condition means based on the approximate formula proposed by Westfall, Kenny, & Judd, (2014) for linear mixed models with contrast codes and single-degree-of-freedom tests (see also Judd, Westfall, & Kenny, 2017). Results of the follow-up analyses will be provided in the text. We do not report degrees of freedom and *p* values. Instead, statistical significance at approximately the 0.05 level is indicated by $|t|$ *or* $|z| \geq 1.96$. (Schotter et al., 2014).

Results

Data for six participants were dropped from the analyses, as the Swanson Sentence Span was terminated incorrectly and, thus, a reliable score for working-memory capacity could not be calculated. In addition, items to which participants had not responded in time on the target or spill-over sentences (i.e., within 10 sec) were excluded from all analyses (resulting in a total loss of 0.4% of the data). Reading times were log-transformed to correct for right-skewness. On the memory task, participants were generally proficient in distinguishing whether they had read the information of a sentence or not. Averaged across all targets, they scored 79% correct (*SD* = 40). On sentences originating from the task (target, context, and neutral) they scored on average 76% correct (*SD* = 42). On distractor sentences they scored on average 90% correct (*SD* = 30). This shows they had read the texts attentively.

Effects of reading goals on validation

On target and spill-over reading times (see Table 5.2 for descriptive statistics) we observed main effects of goal but no interactions with congruency or accuracy (see Table 5.3 and 5.4 for fixed-effects estimates and associated statistics). The main effects of goal indicated that participants that read for study showed longer reading times than participants that read for comprehension, both on target (M_{study} = 2723 ms, SE_{study} = 28; M_{comp} = 2488 ms, SE_{comp} = 25, β = 0.097, *SE* = 0.03, *t* = 2.277, *d* = 0.23) and spill-over sentences (M_{study} = 3294 ms, SE_{study} = 34; M_{comp} = 3027 ms, SE_{study} = 31, β = 0.102, *SE* = 0.05, *t* = 2.138, *d* = 0.24).

On memory scores we observed a main effect of goal (β = 0.289, *SE* = 0.11, *z* = 2.604, *OR* = 1.32) and a goal * congruency interaction (see Table 5.5 for fixed-effects estimates and associated statistics). To interpret the main and interaction effects of congruency and goal, we conducted post-hoc pair-wise comparisons. Overall, congruent targets were remembered better than incongruent targets both when participants read for study (β = 0.561, *SE* = 0.01, *z* = 5.042, *OR* = 1.75) and when they read for comprehension (β = 0.197, *SE* = 0.01, *z* = 2.004, *OR* = 1.22). As displayed in Figure 5.1b this congruency effect was most prominent when participants read for study. This modulation of the effect of congruency emerged because goal had a profound influence on memory for congruent targets, i.e., targets of congruent texts were remembered better when individuals read for study than when they read for comprehension (β = 0.463, *SE* = 0.137, *z* = 3.389, *OR* = 1.59), yet no reliable simple main effect of goal was observed for incongruent targets (β = 0.099, *SE* = 0.13, *z* = 0.771).

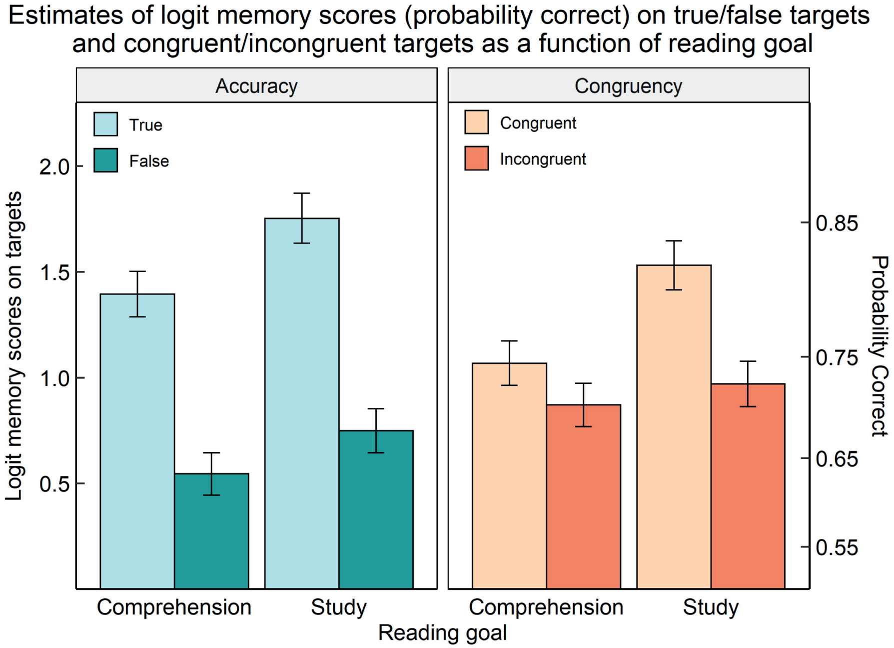

Figure 5.1. Fixed effect estimates of the logit memory performance scores (probability correct) on (a) true and false targets and (b) congruent and incongruent targets as a function of reading goal (comprehension or study). Scales of exponentiated log-values (i.e., approximating untransformed values) are provided as secondary y-axes on the right side of the graphs.

Table 5.2. Mean reading times and standard deviations (in ms) and mean novelty scores and standard deviations at the regions of interest (target and spill-over sentence) for the experimental manipulations (target congruent/incongruent, target true/false and reading goal comprehension/study).

Reading goal	**Accuracy**	**Congruency**	**Target**		**Spill-over**		**Novelty**	
			M	***SD***	***M***	***SD***	***M***	***SD***
Comprehension	True	Congruent	2266	1009	2890	1353	32.10	25.81
		Incongruent	2421	1137	2953	1395	34.52	26.86
	False	Congruent	2523	1206	3095	1572	39.66	26.77
		Incongruent	2747	1323	3174	1530	38.45	26.89
Study	True	Congruent	2440	1126	3172	1505	35.07	24.82
		Incongruent	2646	1176	3213	1418	35.34	24.90
	False	Congruent	2770	1336	3348	1579	39.09	24.40
		Incongruent	3045	1333	3446	1587	38.60	23.74

Table 5.3. Mean memory performance scores on targets (in %) and standard deviations for the experimental manipulations reading goal (comprehension/study), congruency with context (target congruent/incongruent) and accuracy with background knowledge (target true/false).

Reading goal	**Accuracy**	**Congruency**	**Memory performance**	
			M	***SD***
Comprehension	True	Congruent	81	39
		Incongruent	77	42
	False	Congruent	64	48
		Incongruent	61	49
Study	True	Congruent	88	33
		Incongruent	80	40
	False	Congruent	72	45
		Incongruent	61	49

Table 5.4. Fixed effects estimates and the associated statistics of the sum-coded models fitted for log-transformed reading times on target sentences.

Fixed effect	**Beta**	***SE***	***t***
Intercept	7.772	0.031	247.367
Reading goal	0.097	0.043	2.277
Accuracy	0.118	0.011	11.079
Congruency	0.081	0.011	7.616
Novelty	0.001	0.000	3.717
WMC	-0.028	0.029	-0.989
Reading goal * Accuracy	0.039	0.021	1.833
Reading goal * Congruency	0.024	0.021	1.118
Accuracy * Congruency	0.015	0.021	0.716
Reading goal * Novelty	0.000	0.000	0.728
Accuracy * Novelty	-0.001	0.000	-2.437
Congruency * Novelty	-0.000	0.000	-0.667
Reading goal * WMC	0.013	0.057	0.233
Accuracy * WMC	0.005	0.014	0.391
Congruency * WMC	0.015	0.014	1.032
Reading goal * Accuracy * Congruency	-0.003	0.043	-0.063
Reading goal * Accuracy * Novelty	-0.000	0.001	-0.104
Reading goal * Congruency * Novelty	0.000	0.001	0.489
Accuracy * Congruency * Novelty	0.001	0.001	1.773
Reading goal * Accuracy * WMC	0.046	0.028	1.637
Reading goal * Congruency * WMC	0.043	0.028	1.539
Accuracy * Congruency * WMC	-0.027	0.028	-0.950
Reading goal * Accuracy * Congruency * Novelty	-0.000	0.002	-0.255
Reading goal * Accuracy * Congruency * WMC	-0.048	0.056	-0.855

Note. The following R code was used: Reading times ~ 1 + Reading goal * Accuracy * Congruency * Novelty + Reading goal * Accuracy * Congruency * Working Memory + (1|Subject) + (1|Item).

Table 5.5. Fixed effects estimates and the associated statistics of the sum-coded models fitted for log-transformed reading times on spill-over sentences.

Fixed effect	***Beta***	***SE***	***T***
Intercept	7.961	0.035	225.165
Reading goal	0.102	0.048	2.138
Accuracy	0.061	0.010	5.875
Congruency	0.026	0.010	2.506
Novelty	0.001	0.000	3.919
WMC	-0.030	0.032	-0.941
Reading goal * Accuracy	0.000	0.021	0.011
Reading goal * Congruency	-0.005	0.021	-0.222
Accuracy * Congruency	0.012	0.021	0.565
Reading goal * Novelty	0.001	0.000	2.474
Accuracy * Novelty	-0.001	0.000	-1.575
Congruency * Novelty	0.000	0.000	0.550
Reading goal * WMC	-0.007	0.064	-0.117
Accuracy * WMC	-0.004	0.014	-0.286
Congruency * WMC	0.010	0.014	0.721
Reading goal * Accuracy * Congruency	-0.008	0.042	-0.191
Reading goal * Accuracy * Novelty	-0.001	0.001	-1.184
Reading goal * Congruency * Novelty	0.001	0.001	1.610
Accuracy * Congruency * Novelty	-0.000	0.001	-0.479
Reading goal * Accuracy * WMC	0.014	0.027	0.513
Reading goal * Congruency * WMC	-0.012	0.027	-0.444
Accuracy * Congruency * WMC	-0.039	0.028	-1.418
Reading goal * Accuracy * Congruency * Novelty	-0.000	0.002	-0.052
Reading goal * Accuracy * Congruency * WMC	0.114	0.055	2.077

Note. The following R code was used: Reading times ~ 1 + Reading goal * Accuracy * Congruency * Novelty + Reading goal * Accuracy * Congruency * Working Memory + (1|Subject) + (1|Item).

Table 5.6. Fixed effects estimates and the associated statistics of the sum-coded models fitted for memory scores on targets.

Fixed effect	**Beta**	***SE***	***z***
Intercept	1.113	0.077	14.472
Reading goal	0.289	0.111	2.604
Accuracy	-0.953	0.076	-12.493
Congruency	-0.371	0.076	-4.895
Novelty	-0.001	0.002	-0.398
WMC	0.002	0.074	0.030
Reading goal * Accuracy	-0.139	0.152	-0.912
Reading goal * Congruency	-0.341	0.152	-2.246
Accuracy * Congruency	0.144	0.152	0.951
Reading goal * Novelty	-0.002	0.003	-0.809
Accuracy * Novelty	0.006	0.003	1.931
Congruency * Novelty	-0.002	0.003	-0.592
Reading goal * WMC	0.234	0.147	1.590
Accuracy * WMC	-0.102	0.101	-1.012
Congruency * WMC	0.014	0.101	0.143
Reading goal * Accuracy * Congruency	-0.080	0.303	-0.265
Reading goal * Accuracy * Novelty	-0.002	0.006	-0.397
Reading goal * Congruency * Novelty	-0.005	0.006	-0.822
Accuracy * Congruency * Novelty	-0.006	0.006	-1.067
Reading goal * Accuracy * WMC	-0.248	0.200	-1.237
Reading goal * Congruency * WMC	-0.119	0.200	-0.594
Accuracy * Congruency * WMC	-0.001	0.202	-0.007
Reading goal * Accuracy * Congruency * Novelty	0.010	0.012	0.893
Reading goal * Accuracy * Congruency * WMC	-0.184	0.401	-0.459

Note. The following R code was used: Memory Performance ~ 1 + Reading goal * Accuracy * Congruency * Novelty + Reading goal * Accuracy * Congruency * Working Memory + (1|Subject) + (1|Item)

Text-based vs. knowledge-based validation

We observed inconsistency effects of congruency (β = 0.081, *SE* = 0.01, *t* = 7.616, *d* = 0.26) and accuracy (β = 0.118, *SE* = 0.01, *t* = 11.079, *d* = 0.26) on targets, with longer reading times for incongruent (*M* = 2707 ms, *SE* = 27) than congruent (*M* = 2493 ms, *SE* = 25) targets and longer reading times for false (*M* = 2765 ms, *SE* = 28) than true targets (*M* = 2438 ms, *SE* = 24), respectively. In addition, we observed spill-over effects of congruency (β = 0.026, *SE* = 0.01, *t* = 2.506, *d* = 0.06) and accuracy (β = 0.061, *SE* = 0.01, *t* = 5.875, *d* = 0.13), with longer reading times on spill-over sentences following incongruent (*M* = 3190 ms, *SE* = 32) than congruent targets (*M* = 3119 ms, *SE* = 32) and longer reading times on spill-over sentences following false (*M* = 3260 ms, *SE* = 34) than true targets (*M* = 3051 ms, *SE* = 30).

On memory for target information, we observed main effects of accuracy (β = -0.953, *SE* = 0.08, *z* = -12.493, *OR* = 2.53) and congruency (β = -0.371, *SE* = 0.08, *z* = -4.895, *OR* = 1.46): true targets (*M* = 0.81, *SE* = 0.01) were remembered better than false targets (*M* = 0.65, *SE* = 0.01) and congruent targets (*M* = 0.76, *SE* = 0.01) were remembered better than incongruent targets (*M* = 0.70, *SE* = 0.01) (Figure 5.1a and 5.1b).

To investigate whether the effects we observed of congruency and accuracy on readers' memory for target information were mediated by their reading times on targets we conducted multilevel structural equation modeling (MSEM) in Mplus (version 7.31; Muthén & Muthén, 1998-2012). We specified our MSEM consistent with the recommendations of Preacher, Zyphur, & Zhang, (2010) for modeling multilevel mediation when all variables contain both Level 1 (within-person) and Level 2 (between-person) variance (i.e., 1-1-1 mediation). Mediation analysis was performed separately for congruency and accuracy. For both manipulations we tested a 1-1-1 mediation model with a cross-classified structure with random effects for subjects and items that included either congruency (congruent/incongruent) or accuracy (true/false) as a level 1 predictor, (log-transformed) reading times on targets as mediator (level 1) and memory performance scores on targets as dependent variable (level 1). For the estimation, a Bayesian procedure (BAYES estimator in Mplus) was used. Results showed that the within indirect effect of congruency on memory performance through reading times was significant (β = -0.003, *SD* = 0.002, *p* = 0.033, 95% CI [-0.006, 0.000]); Longer reading times on incongruent targets (i.e., a larger online incongruency effect) resulted in poorer memory for those targets (i.e., a larger offline incongruency effect). The within indirect effect of accuracy on memory performance scores was not mediated by target reading times (β = 0.001, *SD* = 0.002, *p* = 0.393, 95% CI [-0.004, 0.005]).

Reader factors and validation

Novelty

We observed main effects of novelty on both target and spill-over reading times: reading times increased when the amount of novel information participants encountered increased. In addition, we observed an accuracy * novelty interaction on target reading times and a goal * novelty interaction on spill-over reading times. We observed no effects of novelty on memory scores.

To interpret these interactions, we conducted post-hoc comparisons by centering the model on the novelty ratings on one standard deviation below (11) and above (62) the mean, respectively. As illustrated in Figure 5.2, comparisons for the target sentences revealed that reading times were generally longer for false targets (*M* = 2765 ms, *SE* = 28) than for true targets (*M* = 2438 ms, *SE* = 24). This inaccuracy effect was most prominent in texts that contained the least novel information for participants (β = 0.14, *SE* = 0.02, *t* = 9.274, *d* = 0.33). The inaccuracy effect diminished as a function of novelty (β = 0.09, *SE* = 0.02, *t* = 5.844, *d* = 0.20). This modulation of accuracy effects emerged because novelty had an influence on the reading times of true targets (i.e., true targets of texts with low novelty scores were read faster than true targets of texts with high novelty scores; β = 0.0014, *SE* = 0.0003, *t* = 4.473), yet no reliable simple main effect of novelty was observed for false targets (β = 0.0004, *SE* = 0.0003, *t* = 1.247).

As illustrated in Figure 5.3, post-hoc pair-wise comparisons for the spill-over sentences revealed that reading times were longer for participants that read for study than for participants that read for comprehension. This effect of goal was more prominent if texts contained more new information for participants (β = 0.14, *SE* = 0.05, *t* = 2.757, *d* = 0.30) and diminished for texts that contained less new information, to such an extent that for texts with lower novelty ratings no reliable differences between reading times were observed for different reading goals (β = 0.08, *SE* = 0.05, *t* = 1.603). This modulation of the effect of goal emerged because participants that read for study showed increased reading times as the amount of novel information increased (β = 0.0015, *SE* = 0.0003, *t* = 4.404), whereas participants that read for comprehension showed no such effect of novelty (β = 0.0004, *SE* = 0.0003, *t* = 1.276).

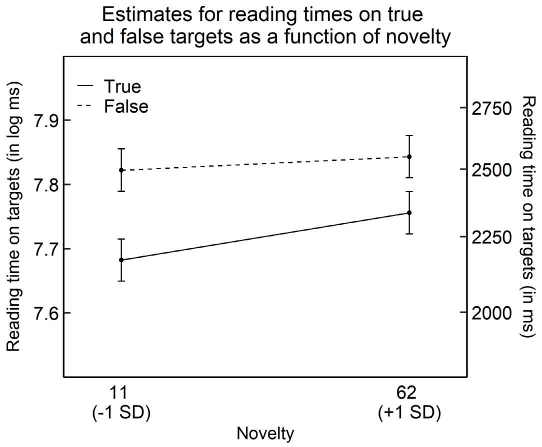

Figure 5.2. Fixed effect estimates for log transformed reading times on true and false targets (in log ms) as a function of novelty (i.e., participants ratings of how much of the information they encountered in the next was novel to them on a scale from 0 (nothing is new) to 100 (everything is new)). Error bars represent the SE of the mean at novelty ratings one standard deviation above (62) and below (11) the mean. Scales of exponentiated log-values (i.e., approximating untransformed values) are provided as secondary y-axes on the right side of the graphs (in ms).

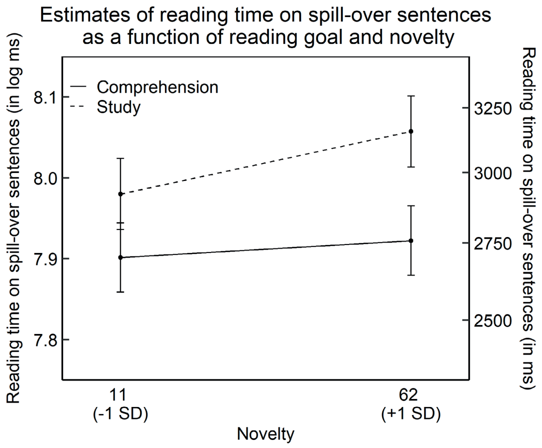

Figure 5.3. Fixed effect estimates of reading times on spill-over sentences (in log ms) for participants that read for study and participants that read for comprehension as a function of novelty (i.e., participants ratings of how much of the information they encountered in the next was novel to them on a scale from 0 (nothing is new) to 100 (everything is new). Error bars represent the SE of the mean at novelty ratings one standard deviation above (62) and below (11) the mean. Scales of exponentiated log-values (i.e., approximating untransformed values) are provided as secondary y-axes on the right side of the graphs.

Working-memory capacity

Overall, we observed no effects of working memory capacity, apart from a four-way interaction between goal, congruency, accuracy and working memory capacity on spill-over sentence reading times (Figure 5.4). To understand this four-way interaction we ran separate linear models (including the full factorial interactions between the fixed factors congruency, accuracy and working memory capacity for participants that read for study and participants that read for comprehension.

We observed a main effect of accuracy in both models: Longer reading times on spill-over sentences following false targets (M_{study} = 3397 ms, SE_{study} = 49; M_{comp} = 3134 ms, SE_{comp} = 46) than on spill-over sentences following true targets (M_{study} = 3193 ms, SE_{study} = 45; M_{comp} = 2922 ms, SE_{comp} = 41) both when participants read for study (β = 0.060, *SE* = 0.02, *t* = 3.914, *d* = 0.13) and when they read for comprehension (β = 0.063, *SE* = 0.01, *t* = 4.630, *d* = 0.14). However, the two models also showed an important difference: In addition to the main effect of accuracy, participants that read for comprehension showed a three-way interaction between accuracy, congruency and working memory capacity (β = -0.089, *SE* = 0.04, *t* = -2.174), whereas participants that read for study showed no other main effects or interaction effects. These results indicate that the four-way interaction in our main model is the result of an accuracy * congruency * working memory capacity interaction that only occurs when participants read for general comprehension, not when they read for study (see Figure 5.4).

To further characterize the three-way interaction in the reading for general comprehension condition, we conducted post-hoc pairwise comparisons separately for lower-capacity readers and higher-capacity readers by centering the model on working-memory scores one standard deviation below (1.5) and above (3.0) the mean (*M* = 2.25, *SD* = 0.75), respectively. In comparison to the spill-over sentences in the true-congruent condition, lower-capacity readers show longer reading times for sentences following false-incongruent targets (β = 0.083, *SE* = 0.03, *t* = 2.885, *d* = 0.20), but not for sentences following true-incongruent targets (β = 0.027, *SE* = 0.03, *t* = 0.951) and false-congruent targets (β = 0.026, *SE* = 0.03, *t* = 0.907) (see Figure 5.4, left side). For higher-capacity readers a different pattern is observed. In comparison to the spill-over sentences in the true-congruent condition, higher-capacity readers show longer reading times for sentences following false-congruent targets (β = 0.082, *SE* = 0.03, *t* = 2.671, *d* = 0.18), true-incongruent targets (β = 0.069, *SE* = 0.03, *t* = 2.230, *d* = 0.15), and false-incongruent targets (β = 0.091, *SE* = 0.03, *t* = 2.969, *d* = 0.20) (see Figure 5.4, right side).

Taken together, results of the post-hoc analyses of the four-way interaction show knowledge inaccuracy spill-over effects in both reading goal conditions. However, for participants that read for comprehension spill-over effects are modulated by the readers' working-memory capacity: Higher-capacity readers show spill-over effects of inaccuracy (world knowledge) and incongruency (contextual), whereas lower-capacity readers show a more restricted pattern with spill-over effects only emerging when target information is inconsistent with both sources.

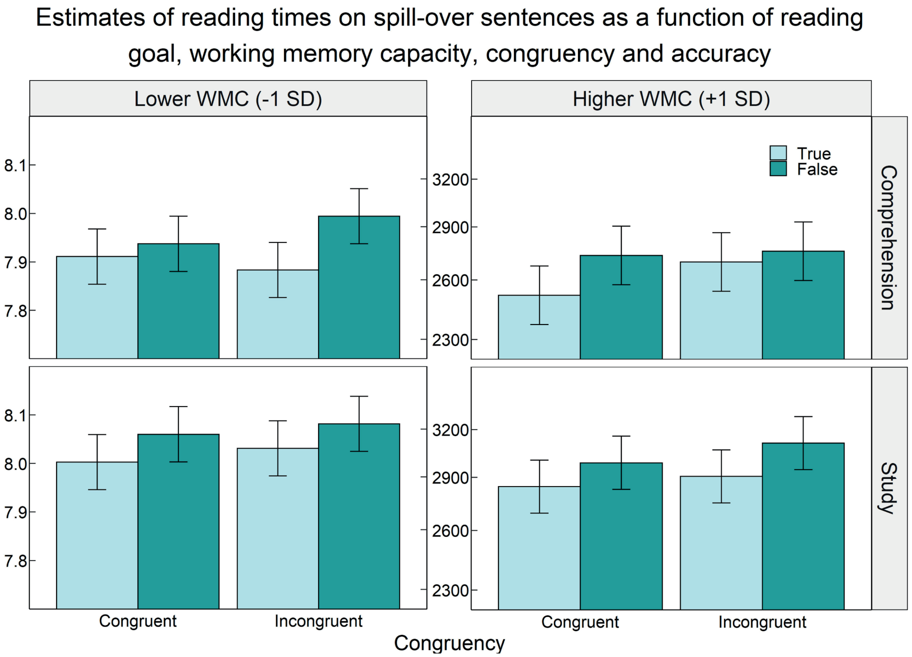

Figure 5.4. Fixed effect estimates of reading times on spill-over sentences (in log ms) for participants with a lower working-memory capacity (one SD below the mean) and participants with a higher working-memory capacity (one SD above the mean) as a function of reading goal (reading for general comprehension/reading for study), congruency (target congruent/incongruent with context) and accuracy (target true/false). Error bars represent the SE of the mean at working memory capacity scores one standard deviation above (2.75) and below (1.5) the mean. Scales of exponentiated log-values (i.e., approximating untransformed values) are provided as secondary y-axes in the center of the graphs.

Discussion

The aim of this study was to determine whether readers' purpose for reading affects online validation processes and readers' memory for (in)consistent information. In addition, we investigated whether and how online validation processes translate into offline (memory) products. In doing so, we distinguished between text-based and knowledge-based validation processes. A secondary aim was to investigate whether these effects were influenced by the novelty of the text information (i.e., the degree to which the topic of the text was novel to each individual reader) and readers' working-memory capacity.

Effect of reading goals on validation

In line with prior findings (Lorch et al., 1993, 1995; van den Broek et al., 2001; Yeari et al., 2015), we observed general effects of reading goal on comprehension, with reading for study resulting in slower reading and in better memory than reading for general comprehension. Furthermore, we found no clear evidence that reading goals influence online validation processes. Reading goals did have distinct effects on readers' offline memory for (in)consistent target information of the texts. Specifically, the observed stronger memory for participants that read to study applied when the targets in the reading task contained information that was congruent with the preceding text but not when they contained incongruent information. In the latter case, reading for study did not result in a stronger memory representation. We elaborate on each of these findings below.

Results showed no evidence that reading goals affect validation processes that occur while readers are processing a text – neither in the target sentences nor in the spill-over sentences. This finding is consistent with the idea that the coherence-detection (or epistemic monitoring) component of validation, as described in the RI-Val model (Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b) and the two-step model of validation (Isberner & Richter, 2014; Richter, 2015; Schroeder et al., 2008), is a passive and routine process that takes place regardless of people's goals for reading a text. Furthermore, we hypothesized that readers may apply a more stringent coherence threshold (a key component of the RI-Val model) when they read for study. In that case, spill-over effects due to incongruent or inaccurate information should be less prominent for people that read for study than for people that read for comprehension, because the former are more inclined to complete the validation process of inconsistent sentences before moving on in the unfolding text. Our results do not support this hypothesis as we did not observe such modulations of spill over as a function of reading goal.

With regard to the repair processes posited by validation models – most explicitly described in the two-step model of validation (Isberner & Richter, 2014; Richter, 2015; Schroeder et al., 2008) – the interpretation of our data is less straightforward. As discussed in the Introduction, elaborative processes to repair and resolve an inconsistent section of a text are under strategic control of the reader and may take place during or after reading a text (Maier & Richter, 2013; Richter, 2011). Our reading time data seems incompatible with the idea that reading goals modulate epistemic elaboration during reading, yet that does not rule out that post-reading repair and reflective processes vary as a function of reading goal – i.e., these validation processes may become more intensive when people read to study.

On a methodological note, it is possible that we did not observe an interaction between reading goals and online validation because sentence-by-sentence reading times are not sensitive enough to detect changes in validation processes elicited by reading goal manipulations. However, sentence-by-sentence reading time measures have been used in other studies investigating the influence of task demands on online validation processes (e.g., Williams et al., 2018). Williams et al. (2018) used changes in task demands (i.e., varying the number of comprehension questions participants had to answer after reading each text) rather than explicit instructions (as in the current study) to manipulate readers' coherence threshold, observing that these subtle changes affected reading times for the target sentences. Thus, sentence reading times in principle are sensitive enough to pick up validation effects. The absence of reading goal effects in the present study therefore suggests that variations in global goals for reading the texts do not (or less strongly) affect validation processes in comparison to properties of the immediate learning context, such as the task demands used by Williams et al. (2018).

Considering the on- and offline results together yields an interesting contrast: reading for study led to more careful processing of all target types; it also led to stronger memory for all textual information *except* for incongruent information which was processed more extensively, just like the other portions of the texts, but in contrast was *not* remembered better. Given that readers did detect all violations – including those involving text incongruencies – as all violations resulted in increased reading times, this pattern suggest that incongruency with the text is dealt with differently than inaccuracy with reader's background knowledge. Because readers that read for study are more likely to put effort into building a comprehensive, coherent representation of the text than are readers with a simple comprehension goal (e.g., Britt et al., 2018; Lorch et al., 1995), they are more likely to try and resolve incongruencies. Indeed, they take more time to read the texts than their counterparts that read for comprehension, thus providing support for this prediction. As noted earlier, this added processing time did not result in better memory for incongruent target information, suggesting that the effort generally did not lead to successful resolution or attained resolution by adjusting the representation of the target information to fit the context (i.e., make it congruent) – and thus lowering memory for the precise target sentence.

Text-based and knowledge-based validation

In addition to the effects of reading goals on validation, the current study considered potential differences between text-based and knowledge-based validation. In line with prior findings (Albrecht & O'Brien, 1993; Menenti et al., 2009; O'Brien et al., 1998, 2004, 2010; O'Brien & Albrecht, 1992; Rapp, 2008; Richter et al., 2009), our results showed inconsistency effects of both text and knowledge violations. Furthermore, the current results replicated those of earlier studies (van Moort et al., 2018, 2021) by showing that knowledge violations generally elicited a prolonged disruption of the reading process frequently spilling over to the next sentence. However, our current results also contradicted prior findings. Unlike in earlier studies, using a similar paradigm, (van Moort et al., 2018, 2020), in the current study we observed prolonged disruptions due to text violations as evidenced by spill-over to the next sentence. These mixed patterns of spill-over effects across studies are puzzling. One possible explanation would be that that subtle variations in samples, instructions, and research methodologies (cf. van Moort et al., 2018; van Moort, Jolles, et al., 2020; van Moort, Koornneef, et al., 2020) affected the settings of readers' coherence thresholds (see RI-Val model; Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b), resulting in small but detectable differences in the amount of spill over across studies.

To obtain a detailed picture of the relation between the on- and offline results for text-based and knowledge-based validation, we conducted a series of mediation analyses. The results showed that the reading times at the target sentence mediated the offline memory results. Specifically, when readers encounter a sentence that is incongruent with the context, the reading time for that sentence increases and the magnitude of this increase, in turn, predicts the decrease in performance on the memory test. Thus, the probability of correctly recalling an incongruent section of a text seems to diminish as the intensity of the repair processes that occur after detecting that incongruency increase. Repair processes may encompass different strategies. For example, readers may adjust the incoming information to make it fit with the representation of the preceding text, they may decide to dismiss the incongruent information and 'remove' it from their developing situation model, and so on. It would be useful for both theory and instruction to investigate the range of repair processes in which readers engage in response to within-text inconsistencies and how the different processes relate to comprehension and memory for the text as a whole.

In contrast, the effects of world knowledge inaccuracies on memory performance were not mediated by reading time. This absence of a mediation effect can be interpreted in several ways. It could mean that reading time disruptions that are observed when readers encounter inaccurate sentences do not index post-detection repair processes. If that is the case, our world knowledge manipulations seem to influence the (offline) memory products primarily via mechanisms that occur after a text has been read. Alternatively, it could mean that the online repair processes to resolve world-knowledge inaccuracies in our materials are relatively straightforward and do not result in detectable changes in sentence reading time. A final possibility is that efforts to repair inaccuracies elicit detectable processing costs but that the amount of time spent on them does not reflect the quality or effectiveness of those processes. In that case increased reading time durations will not correlate with reduced performance on the memory test.

In conclusion, although the mediation analysis cannot tell the full story, it is a powerful tool to decipher whether and how online (reading time) processes translate into offline (memory) products. In the context of our discussion on text-based versus knowledge-based validation, the mediation analyses complement prior findings by indicating that these types of validation have different processing signatures and may trigger different coping mechanisms to protect emerging and final mental representations of readers against inconsistencies (cf. van Moort et al., 2018; van Moort, Jolles, et al., 2020; van Moort, Koornneef, et al., 2020).

Individual variations

We explored whether the above findings were influenced by individual differences; we specifically considered the degree to which the topic of a text was novel to the reader, and reader's working-memory capacity.

With respect to the influence of novelty we found, as predicted, that the processing difference between accurate and inaccurate targets – the amount of conflict a reader experiences – diminished when readers had less knowledge of the topic (i.e., the topic had greater novelty). This finding supports the premise that validation against background knowledge indeed takes places routinely, distinguishing accurate and inaccurate textual information. It also illustrates the importance of topic-relevant or world knowledge in successful comprehension of texts (Alexander & Jetton, 2000; Kintsch, 1988; Myers & O'Brien, 1998; Ozuru, Dempsey, & McNamara, 2009; Samuelstuen & Bråten, 2005; Shapiro, 2004). Interestingly, novelty interacted with accuracy in that increasing novelty resulting in slower reading of accurate but not of inaccurate information. Although one should be cautious with this subtle interaction, one can speculate that it signifies that having knowledge about a topic primarily facilitates textual information that converges with that knowledge, rather than that it hinders processing of conflicting information.

Furthermore, the effect of novelty on spill-over sentence reading times was modulated by reading goal: When the amount of novel information in a text increased, readers tended to slow down on the post-target sentence when they read for study, but not when they read for comprehension. These results suggest that readers engage in deeper or more effortful processing of novel information when the reading goal requires a deep understanding of the text.

With respect to the role of working memory in validation, we considered scenarios in which working-memory capacity would affect the coherence-detection phase and/or the epistemic elaboration phase. Our results did not signal any main effects of differences in working-memory capacity on processing of the target sentences. We did observe an effect of working memory on spill-over sentences as part of a complex (four-way) interaction. When reading for comprehension, the spill-over patterns of higher-capacity readers differed from the spill-over patterns of lower-capacity readers – i.e., arguably more prominent spill-over effects for higher-capacity readers. When reading for study, however, the spill-over patterns for higher- and lower-capacity readers showed no differences. A possible, speculative, explanation for this pattern of results is that when higher-capacity readers are reading for comprehension, they adopt a more lenient processing approach (where processing is allowed to spill-over to the next sentence) than lower-capacity readers. This difference between higher- and lower-capacity readers disappears when people are reading for study: Reading for study may trigger a more stringent processing approach for higher-capacity readers that allows more validation processes to be completed before proceeding to the next sentence. Interpreted as such, these results may also have interesting implications for the coherence threshold of the RI-Val model (Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b), as they suggest that this threshold varies depending on readers' working-memory capacity: Because higher-capacity readers have the capacity to process more information simultaneously they may set a lower coherence threshold than lower-capacity readers, resulting in more 'delayed' processing. The observation that spill-over effects become weaker when higher-capacity individuals read for study also fits this explanation: When reading for study these individuals may set a higher threshold that allows more validation processes to be completed before proceeding to the next sentence. This account cannot provide a perfect explanation for our results, but it raises interesting points for future research.

Conclusions and future directions

The current study investigated whether readers' purpose for reading affects online validation processes and the translation of these processes into the offline mental representation. Results suggest that coherence-detection is a routine aspect of comprehension that is not affected by reading goals (Cook & O'Brien, 2014; Isberner & Richter, 2014a; Singer, 2019; van den Broek & Helder, 2017). The interpretation of the results is less straightforward for the epistemic elaboration component of validation (Isberner & Richter, 2014; Richter, 2015; Schroeder et al., 2008); our results are incompatible with the idea that reading goals modulate the early phases of epistemic elaboration, yet do not rule out that late epistemic phases (including possibly post-text validation processes) are affected by reading goal manipulations. Because reading goals did affect readers' memory for target information, the most parsimonious conclusion is that reading goal influences take place after the initial detection of the inconsistency and also after initial repair processes activated by epistemic elaboration. Determining precisely which after-detection processes are influenced by reading goals and whether the effects we observed are unique to the particular goals for reading used in this study would be fruitful directions for further research. In addition, mapping the time course of such reading goal influences requires more detailed examinations of when and how goals exert their influence (e.g., by assessing the mental representation during reading, immediately after reading a text, and at later points in time).

In addition, reading goal effects depend on the quality of the text, as reading for study improves memory for congruent, but not for incongruent target information. This has important consequences for the interpretation of results from studies investigating the effects of reading purpose, because these studies predominantly use coherent and accurate texts. Moreover, these results raise interesting questions for future research, for example whether an incongruency in a text only affects memory for the incongruent information itself or whether it also affects memory for other (related) elements in the mental representation.

The current results replicate earlier findings that the processes involved in coherence monitoring depend on validation against both contextual information and background knowledge (van Moort, Jolles, et al., 2020; Van Moort et al., 2018; van Moort, Koornneef, et al., 2020). Furthermore, they suggest that reading goals differentially affect processing of text and knowledge violations, respectively, given that reading for study results in longer reading times on both types of violations but only improves memory for the latter. To further examine these differential effects a more detailed overview of these processes is needed. Research methods that have high temporal resolution (e.g., eye tracking, EEG) and research methods that provide more qualitative data (e.g., think-aloud procedures) may be useful in mapping potential differences between text-based and knowledge-based validation. In addition, statistical methods such as mediation analyses can further enlighten us about how online comprehension and validation processes translate into offline memory products. Moreover, to gain insight into the effects of readers' background knowledge (and the extent of this knowledge) on knowledge-based validation processes, future studies could include more extensive assessments of readers' knowledge on a text topic. Finally, the current study focused on text-based and knowledge-based validation processes in the context of reading single texts, but future studies could extend this work by examining when and how readers use these informational sources – and perhaps other informational sources (e.g., readers' prior beliefs; Gilead et al., 2018) – to construct a coherent and adequate mental model when reading multiple texts (e.g., when reading on the web to make an informed decision on a controversial topic; Rouet & Britt, 2011).

We observed minimal effects of working memory on either online processing or offline representation. These results are only partly in agreement with earlier findings from studies using a similar paradigm (van Moort et al., 2020; van Moort et al., 2018). The mixed effects across studies may be attributed to differences between the groups that were tested, or it may illustrate that the role of working memory in validation processes is more complex than initially thought. Including working-memory capacity as a covariate seems insufficient to see which of these possibilities is accurate. Therefore, future studies may include direct manipulations of working memory load during processing (cf. de Bruïne et al., 2021).

There has been a longstanding acknowledgement by reading researchers that one's purpose for reading plays an important role in reading, but a challenge for theories of reading has been to describe when and how reading processes are influenced by reading goals. To deepen our understanding of this issue a more detailed examination of how readers' goals affect component processes of comprehension is needed. Building on the strong tradition of research on goal effects on online comprehension processes and offline products of comprehension, the current study has taken the first step by examining how reading goals affect validation processes. Although reading goals affect readers' general processing, we observed no evidence that they affect the coherence-detection phase of validation. They did influence post-detection processes, differentially affecting readers' memory for incongruent and inaccurate targets. To develop a comprehensive model of reading goal effects, future studies may extend this work by going beyond the impact of reading goals on general comprehension and focus on their effects on specific component processes.