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

4

Differentiating text-based and knowledge-based validation processes during reading: Evidence from eye movements

Abstract

To build a coherent accurate mental representation of a text readers routinely validate information they read against the preceding text and their background knowledge. It is clear that both sources affect processing, but when and how they exert their influence remains unclear. To examine the time course and cognitive architecture of text-based and knowledge-based validation processes we employed eye-tracking methodology. Participants read versions of texts that varied systematically in (in)coherence with prior text or background knowledge. Contradictions with respect to prior text and background knowledge both were found to disrupt reading, but in different ways: The two types of contradiction led to distinct patterns of processes and, importantly, these differences were evident already in early processing stages. Moreover, knowledge-based incoherence triggered more pervasive and longer (repair) processes than did text-based incoherence. Finally, processing of text-based and knowledge-based incoherence was not influenced by readers' working memory capacity.

Keywords: reading comprehension, validation, eye-tracking, background knowledge, (in)coherence

Introduction

Successful comprehension requires readers to build a coherent, meaningful mental representation or situation model of a text (Graesser et al., 1994; van den Broek, 1988; Zwaan & Singer, 2003). An essential aspect of building such mental representation is that readers routinely monitor to what extent incoming information is both coherent and accurate (e.g., Isberner & Richter, 2014a; Singer, 2013). Recent theoretical models of epistemic monitoring (Isberner & Richter, 2014a) and validation (Richter, 2015; Singer, 2013) suggest that such monitoring consists of validation processes involved in detecting possible inconsistencies and by (optional) epistemic elaboration processes involved in attempting to resolve detected inconsistencies (e.g., Isberner & Richter, 2014a; Richter, 2011; Richter et al., 2009; Schroeder et al., 2008).

These monitoring processes can be influenced by various sources of information, most notably contextual information (from preceding text) and the reader's background knowledge. Prior behavioral work using self-paced sentence-by-sentence reading paradigms suggests that each of these two sources has a unique influence on monitoring (van Moort et al., 2018, 2020).

Such results highlight the importance of distinguishing between text-based and knowledge-based monitoring processes, but they do not provide a detailed picture of the constituent processes of validation against each source as reading times aggregate over all such processes. As a result, for example, they cannot distinguish between earlier (e.g., detecting an inconsistency) and later (e.g., repairing an inconsistency) validation processes. Also, the self-paced sentence-by-sentence reading task is somewhat unnatural as readers typically are unable to look back in the text (e.g., Hyönä et al., 2002). The current study aims to provide a more detailed picture of the component validation processes involved in coherence monitoring by adopting eye-tracking methodology. This method offers high temporal resolution and various indices of processing in a more natural reading situation (as texts are presented in their entirety), allowing us to elucidate when and how processes involved in detecting and resolving inconsistencies are influenced by contextual information and background knowledge, respectively.

Building and validating mental representations of texts

Understanding discourse requires comprehension of individual words and sentences as well as integration across sentences, to form a coherent understanding

of the discourse as a whole (Perfetti & Frishkoff, 2008). As readers proceed through a text, they continually use various forms of information – e.g., semantic (the meaning of words), syntactic (grammatical) and pragmatic (their understanding of the world) – to build an overall representation of the discourse meaning (Johnson-Laird, 1983; Kintsch, 1988). To build a coherent and accurate mental representation readers validate incoming information against various sources of information, most notably the preceding text and the readers' background knowledge (Isberner & Richter, 2014a; Nieuwland & Kuperberg, 2008; O'Brien & Cook, 2016a, 2016b; Schroeder et al., 2008; Singer et al., 1992). Successful validation is widely considered to be a prerequisite for comprehension accuracy or, more specifically, situational updating (Cook & Myers, 2004; Ferretti et al., 2008), but theoretical models differ in how they define the validation process.

For example, the RI-Val model of comprehension (Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b) describes validation as one of three processing stages –resonance, integration and validation- that comprise comprehension. According to this 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. Finally, the initial linkages formed by integration are validated against information in memory that is 'readily available' to the reader. Thus, it is validated against information that either already is part of working memory or easily can be made available from long-term memory (e.g., McKoon & Ratcliff, 1995; Myers & O'Brien, 1998; O'Brien & Albrecht, 1992). These contents of active memory includes both portions of the episodic representation of the text (i.e., context) and general world knowledge, therefore each source has the potential to influence validation at any point during comprehension. The three processes –resonance, integration and validation- run parallel but their onset is asynchronous: Activation must produce a minimum of two concepts (or ideas) before integration can begin, and integration must produce a minimum of one linkage before validation can begin. Thus, the RI-Val model presents validation as a single, passive pattern-matching process that is part of comprehension and is involved in detecting mismatches between the linkages made during the integration stage and the contents of active memory. RI-Val focuses on the (in)consistency detection component of coherence monitoring rather than on potential (repair) processes triggered by a detected inconsistency.

A second model describes validation as consisting of two components: (1) epistemic monitoring (i.e., detecting inconsistencies) during a comprehension stage, followed by (2) optional epistemic elaboration processes (i.e., attempting to resolve an inconsistency) during an evaluative stage (e.g., Isberner & Richter, 2014a; Richter, 2011; Richter et al., 2009; Schroeder et al., 2008). According to this model only 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 and carried out routinely and efficiently, i.e., they pose little demands on cognitive resources and are not dependent on readers' processing goals (Richter et al., 2009). However, whereas the RI-Val model focuses in detail on the detection component of coherence monitoring, the two-component model focuses on the resources-demanding processes that may be triggered by the detection of the inconsistency in the epistemic monitoring component. Specifically, readers may initiate evaluation processes, including epistemic elaboration or repair processes, in an attempt to resolve an inconsistency. For example, they may doubt the validity of their current mental representation/situation model (e.g., perhaps they misunderstood earlier parts of the text), they may disbelieve the target sentence rather than their situation model (e.g., perhaps the target sentence contains a mistake), or they may try to solve the comprehension problem by elaborating possible solutions to the apparent inconsistency (Hyönä et al., 2003). These processes are optional and only occur when readers are motivated and have enough cognitive resources available, as they are assumed to be slow, resource-demanding and under at least some strategic control of the reader (Isberner & Richter, 2014a; Richter, 2003, 2015)

These examples illustrate that current theoretical models presume a rudimentary cognitive architecture and time course for validation processes. They generally agree that incoming information is routinely validated against elements of the current situation model or world knowledge during the comprehension stage and that contextual information and background knowledge both have the potential to influence processing (e.g., Cook & Myers, 2004; Kintsch, 1988; O'Brien & Cook, 2016a; Richter, 2011; Richter et al., 2009; Rizzella & O'Brien, 2002; Schroeder et al., 2008; Singer, 2013; van den Broek & Helder, 2017). However, it is unclear when and how contextual information and the readers' background knowledge exert their influence.

General models of discourse comprehension (i.e., without a specific focus on the validation aspect of comprehension) present different viewpoints about the respective influences of context and background knowledge on comprehension. Some accounts presume a fully interactive architecture with contextual information and background knowledge immediately influencing processing (e.g., the memory-based text processing view; Cook et al., 1998b; Gerrig & McKoon, 1998; Myers & O'Brien, 1998; O'Brien & Myers, 1999; Rizzella & O'Brien, 2002). Other accounts presume an architecture in which one of the informational sources plays a more dominant (and sometimes earlier) role. In some of these accounts, background knowledge is regarded as the dominant source (i.e., driving comprehension) as incoming information is first connected to general world knowledge and only later integrated in the discourse context (e.g., Garrod & Terras, 2000; Kintsch, 1988; Sanford & Garrod, 1989). In other of these accounts, contextual information is

4

regarded as the dominant source as it can influence language comprehension immediately (e.g., Hess et al., 1995; Marslen-Wilson & Tyler, 1980; van Berkum et al., 1999) and can fully override background knowledge (e.g., Nieuwland & Van Berkum, 2006).

Within validation research there are a considerable number of empirical investigations of the effects of contextual information and background knowledge on validation processes (e.g., 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). Usually the focus of each investigation is on one source of potential inconsistencies or the other, whereas in reality both sources operate in tandem. With respect to investigations of text-based monitoring detection of within-text incongruencies inevitably depends on background knowledge as well. For example, in paradigms in which targets (e.g., children are building a snowman) presumably are incongruent with preceding context (e.g., it was a hot, sunny day) detection only occurs if readers have certain background knowledge (e.g., snow melts on a hot sunny day). Even blatant incongruencies (e.g., a car is described as solid blue in one sentence, but as solid red in the next) require at least a minimal amount of background knowledge (e.g., red and blue are colors and something cannot be solid red and blue at the same time). Thus, although the role of background knowledge is implied because it is essential for detecting the (in)congruency of textual targets it is not explicitly included as a factor. With respect to studies that do include both contextual and world knowledge manipulations the central question tends to be whether context can override (erroneous) world knowledge, not whether text-based monitoring is an independent process (e.g., Creer et al., 2018; Menenti et al., 2009; Walsh et al., 2018). As a result, it is difficult to distinguish between the respective impacts of textual information and background knowledge and to define possibly unique influences.

To address this issue, van Moort et al. (2018) contrasted validation against background knowledge and validation against prior text in a single design. Participants read expository texts about well-known historical topics in a self-paced, sentence-by-sentence manner. Each text contained a target sentence that was either true (e.g., the Statue of Liberty was delivered to the United States) or false (e.g., the Statue of Liberty was not delivered to the United States) relative to the reader's background knowledge and that was either supported or called into question by the preceding context (e.g., context that described that the construction of the statue went according to plan vs. context that described problems that occurred during construction of the statue). Results indicated that both prior text and background knowledge influenced readers' moment-by-moment processing on targets, but only inaccuracies with background knowledge elicited spill-over effects. This suggests that both sources of information have unique influences on processing. Furthermore, a recent study used the same reading paradigm while collecting neuroimaging data

(fMRI) to examine the neural underpinnings of text-based and knowledge-based validation (van Moort et al., 2020). Consistent with Van Moort et al. (2018), the neuroimaging data suggested a 'division of labor' for text-based and knowledge-based validation processes. The medial Prefrontal Cortex (mPFC) seems to be oriented towards knowledge-based processing, whereas the right Inferior Frontal Gyrus (right IFG) is more involved in text-based processing. Interestingly, the Precuneus and the left Inferior Frontal Gyrus (left IFG) seem to combine the information provided by a text with the information stored in long-term memory. Taken together, the results from these two studies suggest that both sources impact processing and that text-based and knowledge-based validation processes may involve (partially) different cognitive mechanisms.

The current study

The aim of this study was to provide insight into component validation processes involved in coherence monitoring and into when and how contextual information and background knowledge influence these processes. Specific questions were whether (parts of) the validation process are more knowledge-driven or text-driven, and to what extent text-based and knowledge-based validation processes take place independently or interactively. We employed eye-tracking methodology because it offers the possibility to distinguish between relatively early processing (e.g., first-pass reading times) and later processing (e.g., second-pass reading times) (Cook & Wei, 2017; Rayner, 1998). It also allows for relative naturalistic reading as texts are presented in their entirety (e.g., Hyönä et al., 2002). The basic assumption of eye tracking methods is that increased processing demands, e.g., when readers encounter a comprehension problem in a text, are associated with increased processing time or changes in the pattern of fixations (e.g., Frazier & Rayner, 1982; Hyönä et al., 2003; Rayner et al., 2004; Rayner & Slattery, 2009; Rinck et al., 2003; Stewart et al., 2004). Such changes are assumed to be indicative of underlying processes. For example, readers may detect and attempt to resolve incoherence by spending more time on the critical regions (Yuill & Oakhill, 1991), by engaging in rereading activities to look for the possible source of the incoherence (Hyönä et al., 2003; Zabrucky & Ratner, 1986), or by making regressions to earlier parts of the text to reinstate information from the text that they would like to elaborate or reactivate in working memory (Hyönä & Lorch, 2004).

Participants read expository texts containing information that conflicted with the preceding text and/or readers' background knowledge (based on Van Moort et al., 2018), while their eye movements were recorded as they freely read through the texts. Assuming that initial validation processes (i.e., detection of inconsistencies)

occur relatively early in processing and elaboration and repair processes (i.e., attempts to resolve the inconsistency) occur later in processing, recording eye movements allows us to investigate whether text and background knowledge affect early and later validation processes independently or interactively and, conversely, whether early and late processes (or both) are predominantly knowledge-driven or context-driven.

The secondary aim was to investigate whether working memory capacity affects text-based and knowledge-based validation processes differentially and whether validation components are impacted by individual differences in processing capacity. An important assumption of most models of validation is that incoming information can only be validated against information that is available and activated during comprehension. Therefore, working memory capacity is likely to play a role in validation as it limits the amount of information that can be activated (e.g., Hannon & Daneman, 2001; Singer, 2006). For instance, two-component models of validation suggest that epistemic elaboration or repair processes may be particularly impacted by individual differences in processing capacity, as they are assumed to be reader-initiated and resource-demanding. Working memory indeed has been found to play a role in at least knowledge-based validation (van Moort et al., 2018) but it is unclear whether it impacts epistemic elaboration as well. To investigate these possibilities we obtained a measure of participants' working memory capacity.

Method

Participants

Forty-seven native speakers of Dutch (39 females, 8 males) aged 18-27 years (M = 21, SD = 2) participated in this study[11]. All participants had normal or corrected-to-normal eyesight and none had diagnosed reading or learning disability. Participants provided written informed consent prior to testing and were paid 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.

---

[11] In total 70 participants were tested. Due to technical issues with the eye tracker at the start of the study the data of 23 participants was of insufficient quality and could not be analyzed.

Materials

We used the texts of Chapter 2 of this dissertation (based on Rapp, 2008). The texts are about well-known historical topics. The texts were normed to ensure that the presented facts were common knowledge in our sample (see Chapter 2 of this dissertation for a more detailed description of the norming study). Each text contains a target sentence that is either true or false (with respect to the readers' background knowledge); at the same time the preceding text could either support or call into question the information in the targets. More specifically, the context could bias towards either the true or the false target, making the context either congruent or incongruent with the target (see sample text in Table 4.1). Four different versions of each of the 80 texts were constructed, by orthogonally varying the target sentence (i.e., true versus false) and the context prior to the target sentence (i.e., congruent versus incongruent with target). 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.

Each text consisted of ten sentences (see Table 4.1 for a sample text). Sentence 1-2 were identical among all conditions, providing an introduction to the topic. Sentences 3–7 differed in content, depending on context condition (congruent/incongruent). On average, the context biasing towards the true target consisted of 64 words (SD = 4.20) and 399 characters (SD = 23.48) and the context biasing towards the false target consisted of 66 words (SD = 4.30) and 407 characters (SD = 22.79). Sentence 8 was the target sentence and was either true or false. Overall, targets were equated for length: both true (SD = 1.92) and false (SD = 1.90) targets contained on average 9 words and both true (SD = 10.51) and false (SD = 10.42) targets contained on average 60 characters (including spaces and punctuation). Half of the true targets and false targets included the word 'not' or 'never' (e.g., "Jack the Ripper was never caught and punished for his crimes."), and half did not (e.g., "The Titanic withstood the damage from the iceberg collision."). Sentence 9-10 concluded the text. They contained a general conclusion that did not elaborate on the fact potentially called into question in the target sentence and maintained historical accuracy. On average, the texts contained 121 words (SD = 5.66) and 766 characters per text (SD = 37.63), 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 target (true or false) and text context (congruent or incongruent with target) on each list. Each list was randomized. Each participant was assigned to one list and, hence, read one version of each text.

Table 4.1. Sample text with the four text versions (translated from Dutch original)

	Knowledge accuracy	
	Target true	**Target false**
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.

(Left margin label spanning the full table: **Text congruency**)

Reading task

Participants read 80 texts while eye movements were recorded. The texts were presented as a whole and participants were instructed to read for comprehension at their normal pace and to advance to the next text by pressing a button. A fixation cross was presented at the position of the first word of the first sentence in between texts for 300ms. The task consisted of two blocks of 40 texts. Each block started with a calibration of the eye tracking apparatus. Participants performed a short practice block.

Apparatus

Eye movements were recorded using an EyeLink 1000 desktop mounted eye tracker of the SR Research Company (Oakville, Canada). Sampling frequency was 1000 Hz, and spatial accuracy was approximately 0.4°. Viewing was binocular, but only the right eye was tracked. A chin-and-head rest was used to minimize participants' head movements. The texts were presented in their entirety on a 19-inch screen at approximately 65 cm from the participant.

Measures

Working memory capacity

Working memory capacity was measured with 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 while they progress in the test. 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 the participant's error rate exceeds a given threshold. Participants earn points for each correct answer on the comprehension questions and each correctly recalled set of words. The sum of these points is the index of working memory capacity.

Procedure

Participants were tested individually. The eye tracker was calibrated by means of a 9-point calibration grid that covered the entire computer screen. Upon successful calibration participants completed the reading task. Next, they completed

the Swanson Sentence Span Task and a questionnaire assessing their background knowledge on the text topics. The duration of the total session was approximately 90 minutes.

Eye-fixation measures

For each text eye-fixation measures for two regions of interest were calculated: the target sentence and the spill-over sentence (the sentence following the target). We examined several measures for each of these regions. First-pass reading times reflects initial processing of a sentence (Rayner et al., 1989) and were computed by summing all first-pass fixations for each word (all fixations on a word from the first fixation on that word until the first time the reader exits that word) in the sentence. First-pass probability of a regression reflects the probability that a regression was made to an earlier section of the text and was computed for each word in a sentence (Clifton Jr et al., 2007). Second-pass reading times (or re-reading times) reflect later processing or re-processing of the words in a sentence after the words were exited for the first time (Rayner et al., 1989). They were computed by summing all fixations on each word in the sentence excluding first-pass fixations on these words. The probability of rereading reflects the probability that the sentence was read again after the reader exited the sentence for the first time and was computed binarily (rereading present vs. absent) for each sentence. The regression path duration or go-past duration is computed by summing all fixations from first entering a region until exiting in the forward direction (Duffy et al., 1988; Rayner & Duffy, 1986; Rayner & Liversedge, 2012). This measure includes any regression out of a region prior to moving forward in the text.

Analyses

To investigate the effects of the manipulations on the reading process we conducted mixed-effects linear regression analyses using the R package LME4 version 1.1.21 (D. Bates et al., 2015). For each measure on each sentence we started with a full interactional model that included the interaction between the fixed factors background knowledge (target true/false), text (target congruent/incongruent), working memory capacity (median centered), and the random factors subjects and items. Effect coding was applied in the main analyses (true was coded as -0.5 and false as 0.5; congruent was coded as -0.5 and incongruent as 0.5). We did not include random slopes in our models. 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 4.3 and 4.4). To obtain

fixed-effects estimates and the associated statistics for the relevant simple effects of an interaction, we performed pairwise comparisons. The results of the follow-up analysis are provided in the text. As it is not clear how to determine the degrees of freedom for the t statistics estimated by mixed models for continuous dependent variables (Baayen, 2008), we do not report degrees of freedom and *p* values. Instead, statistical significance at approximately the 0.05 level is indicated by $|t| \geq 1.96$. (e.g., Schotter et al., 2014).

Results

Data for one experimental text was dropped from the analyses, as it concerned Stephen Hawking whose death changed the truth value of the text. For the regions of interest (target and spill-over sentence) in the other texts first-pass reading time, first-pass probability of a regression, re-reading probability, second-pass reading time, and regression path duration were determined (see Table 4.2 for descriptive statistics).

Target sentence

First-pass reading times

Wald chi-square tests revealed a main effect of background knowledge, but not of text. Furthermore, we observed a background knowledge * text interaction on the log-transformed first-pass reading times (see Table 4.3 for fixed-effects estimates and associated statistics). Post-hoc multiple comparisons showed increased first-pass reading times for false targets than true targets, both when the target was presented in a congruent context (β = -0.07, *SE* = 0.02, *z* = -3.92) and when it was presented in an incongruent context (β = -0.12, *SE* = 0.02, *z* = -6.90). This effect of background knowledge was modulated by congruency of the target with the preceding context: False targets presented in a incongruent context (e.g., target states that the Statue of Liberty was not delivered to the United States but context suggests it was) elicited longer first-pass reading times than false targets presented in a congruent context (e.g., the target states that the Statue of Liberty was not delivered to the United States and context also suggests that it was not) (β = -0.05, *SE* = 0.02, *z* = -2.64). However, true targets showed no effect of (in)congruency with the preceding context (β = 0.01, *SE* = 0.02, *z* = 0.38) (Figure 4.1a). We did not find any effects of working memory capacity on the first-pass reading times (Table 4.3).

4

Table 4.2. Means and standard deviations for the dependent variables at the regions of interest (target and spill-over sentence) for the experimental manipulations regarding text (target congruent / incongruent with context) and background knowledge (target true / false).

Background knowledge		Text	Target		Spill-over	
			M	SD	M	SD
First-pass reading time (in ms)	True	Congruent	1471	737	1897	846
		Incongruent	1454	673	1948	943
	False	Congruent	1611	848	1991	886
		Incongruent	1658	805	2028	920
Second-pass reading time (in ms)	True	Congruent	405	396	476	447
		Incongruent	439	412	496	488
	False	Congruent	480	428	452	457
		Incongruent	543	520	527	559
First-pass probability of regression	True	Congruent	0.15	0.36	0.15	0.35
		Incongruent	0.18	0.38	0.16	0.37
	False	Congruent	0.19	0.39	0.16	0.37
		Incongruent	0.20	0.40	0.17	0.37
Regression path duration (in ms)	True	Congruent	2032	1883	2429	1698
		Incongruent	2134	2714	2701	2687
	False	Congruent	2218	2015	2600	1646
		Incongruent	2360	1619	2773	1870
Re-reading probability	True	Congruent	0.48	0.50	0.50	0.50
		Incongruent	0.54	0.50	0.52	0.50
	False	Congruent	0.58	0.49	0.54	0.50
		Incongruent	0.62	0.48	0.54	0.50

First-pass probability of regression

In addition to a text * background knowledge interaction, results showed main effects of background knowledge and of text on the first-pass probability of regressions (see Table 3 for fixed-effects estimates and associated statistics). Post-hoc multiple comparisons showed that readers were more likely to make regressions on false targets than true targets, both when the target was presented in a congruent context (β = -0.03, SE = 0.06, z = -5.02) and when it was presented in an incongruent context (β = -0.14, SE = 0.06, z = -2.44) (Figure 4.1b). However, there was only an effect of (in)congruency with the context on targets containing true world

knowledge information: Readers were more likely to make first-pass regressions when true targets were presented in an incongruent context than when they were presented in a congruent context (β = -0.02, *SE* = 0.06, *z* = -3.96). False targets showed no effect of (in)congruency with the preceding context (β = -0.08, *SE* = 0.06, *z* = -1.40) (Figure 4.1b). We did not find any effects of working memory capacity on the first-pass probability of a regression (Table 4.3).

Regression path duration

Results showed main effects of background knowledge, text and working memory capacity on the regression path duration (see Table 4.3 for fixed-effects estimates and associated statistics). False targets elicited longer regressions than true targets (Figure 4.1e). Similarly, incongruent targets elicited longer regressions than congruent targets (Figure 4.1e). Furthermore, participants with a larger working memory made shorter regressions than participants with a smaller working memory.

Re-reading probability

Results showed main effects of background knowledge, text and working memory capacity on the probability of rereading targets (see Table 4.3 for fixed-effects estimates and associated statistics). Participants were more likely to reread false targets than true targets (Figure 4.1c). Similarly, they were more likely to reread incongruent targets than congruent targets (Figure 4.1c). Furthermore, participants with a larger working memory were less likely to reread targets than participants with a smaller working memory.

Second-pass reading times

Results showed main effects of background knowledge, text, and working memory capacity on the second pass reading times on targets[12] (see Table 4.3 for fixed-effects estimates and associated statistics). Participants spent more time re-reading false targets than true targets (Figure 4.1d). Similarly, participants spent more time re-reading incongruent targets than congruent targets (Figure 4.1d). Furthermore, participants with a larger working memory spend less time re-reading targets than participants with a smaller working memory.

[12] Note that second-pass reading times were only included in this measure if a second pass was made.

a. Mean first-pass reading time

b. Mean first-pass probability of regression

c. Mean probability of rereading

d. Mean second-pass reading time

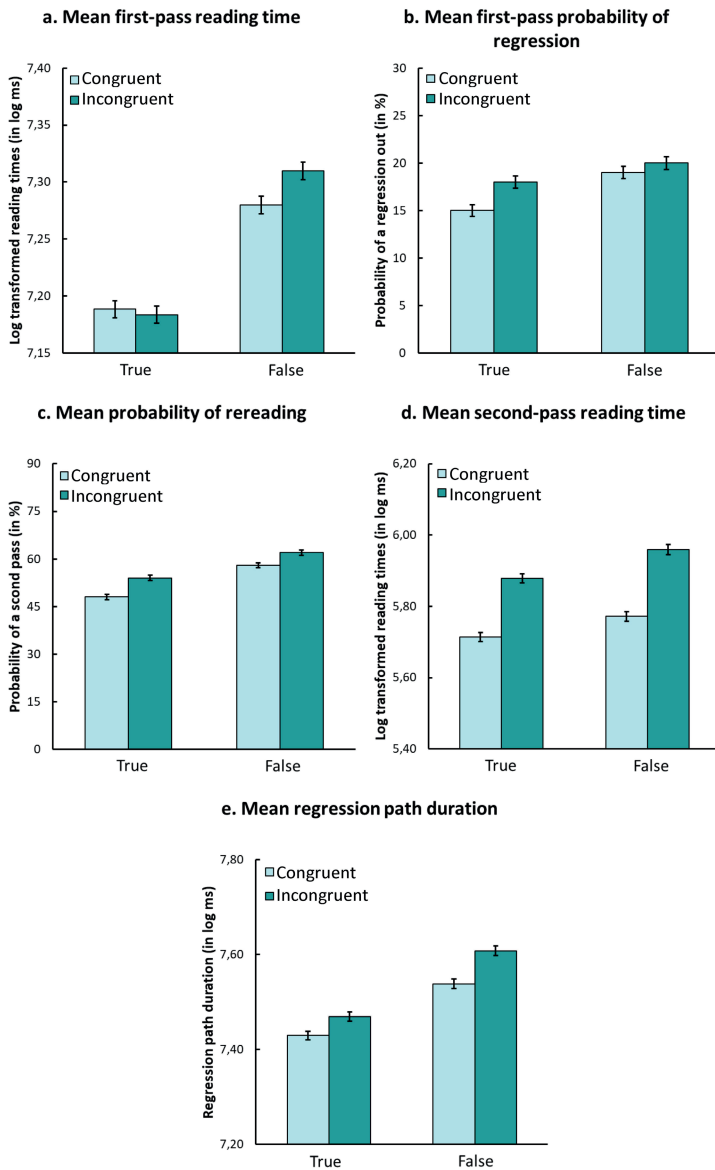**e. Mean regression path duration**

Figure 4.1. Reading patterns for (a) first-pass reading time, (b) first pass probability of regression, (c) probability of rereading, (d) second-pass reading time and (e) regression path duration on target sentences as a function of match with text (congruent / incongruent) and background knowledge (true/false). Error bars represent standard errors of the means.

Table 4.3. Fixed effects estimates and the associated statistics of the sum-coded models fitted for the dependent variables on the target sentence. Note. The following R code was used for all models: dependent variable ~ 1 + Text * Background * Working Memory + (1|Subject) + (1|Item).

Measure	Fixed effect	Statistics		
		B	SE	t/z
First-pass reading time	Intercept	7.23	0.04	197.69*
	Text	0.21	0.01	1.71
	Background	0.10	0.01	7.45*
	WM	-0.05	0.05	-1.64*
	Text * Background	0.05	0.03	2.16*
	Text * WM	0.03	0.02	1.47
	Background * WM	-0.02	0.02	-1.25
	Text * Background * WM	0.02	0.04	0.38
First-pass probability of a regression	Intercept	-1.59	0.09	-18.22*
	Text	0.16	0.04	3.82*
	Background	0.22	0.04	5.20*
	WM	-0.23	0.13	-1.83
	Text * Background	-0.17	0.08	-2.01*
	Text * WM	-0.01	0.07	-0.12
	Background * WM	-0.05	0.07	-0.73
	Text * Background * WM	-0.16	0.13	-1.19
Regression path duration	Intercept	7.50	0.04	192.39*
	Text	0.06	0.02	3.68*
	Background	0.10	0.02	6.33*
	WM	-0.11	0.05	-2.14*
	Text * Background	0.05	0.03	1.43
	Text * WM	0.004	0.02	0.19
	Background * WM	-0.01	0.02	-0.39
	Text * Background * WM	-0.02	0.05	-0.47
Re-reading probability	Intercept	0.29	0.15	2.03*
	Text	0.29	0.08	3.75*
	Background	0.41	0.08	5.23*
	WM	-0.45	0.21	-2.19*
	Text * Background	0.01	0.15	0.09
	Text * WM	0.004	0.12	0.04
	Background * WM	0.01	0.12	0.07
	Text * Background * WM	-0.17	0.23	-0.73
Second-pass reading time	Intercept	5.78	0.04	146.30*
	Text	0.16	0.03	4.60*
	Background	0.07	0.03	2.18*
	WM	-0.13	0.06	-2.23*
	Text * Background	0.04	0.07	0.51
	Text * WM	-0.05	0.06	-0.84
	Background * WM	-0.01	0.06	-0.22
	Text * Background * WM	0.03	0.11	0.30

* indicates |z| or |t| score < 1.96 and thus significance at the 0.05 level.

4

Table 4.4. Fixed-effects estimates and the associated statistics of the sum-coded models fitted for the dependent variables on the spill-over sentence. The following R code was used for all models: dependent variable ~ 1 + Text * Background * Working Memory + (1|Subject)+(1|Item).

Measure	Fixed effect	Statistics		
		B	SE	t/z
First-pass reading time	Intercept	7.48	0.04	183.26*
	Text	0.15	0.01	1.27
	Background	0.04	0.01	3.18*
	WM	-0.02	0.05	-0.45
	Text * Background	0.02	0.03	0.74
	Text * WM	0.02	0.05	1.15
	Background * WM	-0.03	0.05	-1.71
	Text * Background * WM	0.06	0.04	1.75
First-pass probability of a regression	Intercept	-1.77	0.09	-19.19*
	Text	0.18	0.04	2.05*
	Background	0.11	0.04	2.91*
	WM	-0.32	0.14	-2.35
	Text * Background	0.01	0.08	0.18
	Text * WM	-0.05	0.06	-0.81
	Background * WM	0.02	0.06	0.27
	Text * Background * WM	0.09	0.12	0.80
Regression path duration	Intercept	7.71	0.04	175.51*
	Text	0.53	0.01	3.70*
	Background	0.05	0.01	3.70*
	WM	-0.10	0.06	-1.68
	Text * Background	0.01	0.03	0.27
	Text * WM	0.01	0.02	0.23
	Background * WM	-0.004	0.02	-0.18
	Text * Background * WM	0.06	0.04	1.27
Re-reading probability	Intercept	0.13	0.14	0.89
	Text	0.04	0.07	0.49
	Background	0.16	0.08	2.17*
	WM	-0.42	0.20	-2.06*
	Text * Background	-0.08	0.15	-0.52
	Text * WM	-0.10	0.11	-0.89
	Background * WM	-0.06	0.11	-0.48
	Text * Background * WM	0.09	0.23	0.37
Second-pass reading time	Intercept	5.79	0.04	139.91*
	Text	0.002	0.03	-0.06
	Background	0.10	0.03	2.89*
	WM	-0.19	0.06	-3.24*
	Text * Background	0.11	0.07	1.65
	Text * WM	0.01	0.05	0.16
	Background * WM	-0.02	0.05	-0.37
	Text * Background * WM	0.03	0.11	0.31

* indicates |z| or |t| score < 1.96 and thus significance at the 0.05 level.

Spill-over sentence

First-pass reading times

Results showed a main effect of background knowledge on the first-pass reading times on spill-over sentences (see Table 4.4 for fixed-effects estimates and associated statistics). Readers were slower to read spill-over sentences following false targets than those following true targets. We did not find any effects of text and working memory capacity.

First-pass probability of regression

Results showed a main effect of background knowledge, text, and working memory capacity on the first pass chance of regressions on spill-over sentences (see Table 4.4 for fixed-effects estimates and associated statistics). Spill-over sentences that were preceded by false targets were more likely to elicit regressions to earlier text than those preceded by true targets. Similarly, spill-over sentences following incongruent targets were more likely to elicit regressions than those following congruent targets. Furthermore, readers with a larger working memory were less likely to engage in regressions on the spill-over sentence than readers with a smaller working memory.

Regression path duration

Results showed main effects of background knowledge and text on the regression path duration on spill-over sentences (see Table 4.4 for fixed-effects estimates and associated statistics). Spill-over sentences following false targets elicited longer regressions than those following true targets. Similarly, spill-over sentences following incongruent targets elicited longer regressions than those following congruent targets.

Re-reading probability

Results showed main effects of background knowledge and working memory capacity on the re-reading probability on spill-over sentences (see Table 4.4 for fixed-effects estimates and associated statistics). Participants were more likely to reread spill-over sentences preceded by false targets than those preceded by true targets. Furthermore, readers with a smaller working memory were more likely to reread spill-over sentences than readers with a larger working memory.

Second-pass reading times

Results showed main effects of text and working memory capacity on the second pass reading times on spill-over sentences (see Table 4.4 for fixed-effects estimates and associated statistics). Participants spend more time re-reading spill-over sentences following incongruent targets than those following congruent targets. Furthermore, readers with a larger working memory spend less time re-reading the spill-over sentence than readers with a smaller working memory.

Discussion

The aim of the current eye-movement study was to investigate when and how contextual information and background knowledge, respectively, influence validation processes during text comprehension. Additionally, we examined the role of working-memory capacity in validating against these two sources of information.

During first-pass reading, target sentences that contained world-knowledge inconsistencies induced longer reading times than did target sentences that did not contain such inconsistencies. Sentences that contained text incongruencies induced longer reading times than did sentences that contained congruent targets but, interestingly, only when the target contained false (inaccurate) world-knowledge information. In addition, knowledge inaccuracies elicited more regressions during first-pass reading. Text incongruencies also elicited more regressions during first-pass reading, but only in the absence of a knowledge inaccuracy (i.e., when targets contained true world-knowledge information). No interactions between text-based and knowledge-based processing were observed in later processing measures (i.e., regression-path duration, re-reading probability, and second-pass reading time). Instead, these later measures revealed reliable main effects of both context and background knowledge: Readers were more likely to re-read the target and displayed both longer regressions and longer re-reading when they encountered targets that either were false with their world knowledge or incongruent with the preceding context.

In addition to these effects on target sentences, we observed a spill-over effect of knowledge inaccuracies during first pass reading (i.e., longer first-pass reading times on the spill-over sentence if it was preceded by a target containing false world knowledge information) but not of contextual incongruencies. Furthermore, readers were more likely to regress to earlier parts of the text and displayed longer regressions on spill-over sentences following false or incongruent targets. Finally, they were more likely to re-read spill-over sentences following false or incongruent targets and, when they reread a sentence, spent more time doing so.

Working memory did not affect early processing but it did influence later processing (e.g., regressions and re-reading). Readers with a larger working memory were less likely to make regressions or reread targets and, if they do reread, spent less time doing so. Importantly, working memory did not interact with the two types of inconsistencies, indicating that the effects of incongruency with text or inaccuracy with background knowledge did not depend on readers' working-memory capacity.

Differentiating text-based and knowledge-based validation processes

These findings point to several conclusions. First, they support the notion that both incongruencies with context and inaccuracies with background knowledge have a profound impact on validation processes during reading, as evidenced by early and later eye-movements on both target and spill-over sentences. Second, although contextual incongruencies and knowledge inaccuracies are not completely independent (i.e., contextual incongruencies must involve, at the very least, some violation of logic that exist in the reader's background knowledge), the results show that they trigger distinct processes. Third, the distinct patterns of processing for text and knowledge violations already are evident at an early stage in the processing of incoming text information.

Fourth, first-pass reading times and regression probabilities revealed interaction patterns that further differentiate processing of text incongruencies and knowledge inaccuracies. Knowledge inaccuracies consistently disrupted initial processing (i.e., longer first-pass reading times and more first-pass regressions); Text incongruencies also disrupted initial processing but the way in which they did depended on whether there also was a knowledge inaccuracy present. If the target also violated background knowledge, then the contextual incongruency resulted in further slowdown in reading (additional first-pass reading times); In contrast, if the target was accurate according to background knowledge then the contextual incongruency elicited more re-readings (more first-pass regressions). Thus, the target's accuracy/inaccuracy with respect to background knowledge modulates the type of effect that a text incongruence elicits: if incoming text information is inconsistent with both earlier text and the reader's knowledge, then reading becomes extra slow; if the incoming information is inconsistent only with earlier text then it is more likely to be reread.

Fifth, text incongruencies and knowledge inaccuracies differed in the strength of the disruption they caused: Knowledge inaccuracies appeared to induce a more intensive, prolonged disruption of the reading process than did text incongruencies, as reflected in spill-over effects (i.e., effects on first-pass reading of

4

the spill-over sentence) for background-knowledge but not for context contradictions. This pattern is consistent with the notion of dissociable text-based and knowledge-based validation processes.

It is interesting to speculate about possible mechanisms underlying these findings. Drawing an analogy to sentence-processing literature, one possibility is to assume a serial mechanism where knowledge-based information is processed first and text-based information comes into play at a later stage during processing (e.g., Frazier, 1987; Frazier & Fodor, 1978; Frazier & Rayner, 1982). Another possibility is an interactive, constraint-based mechanism in which both sources of information are processed simultaneously (similar to interactive constraint models; Bates & MacWhinney, 1989; MacDonald et al., 1994; Marslen-Wilson, 1973; Tanenhaus & Trueswell, 1995). The serial account fits with the finding that contextual incongruence increases first-pass reading time only if a knowledge-based inaccuracy is detected. However, in the absence of a knowledge inaccuracy, contextual incongruencies do increase the probability of a first-pass regression. In so far as first-pass regressions reflect early processes, this latter finding suggests an interactive mechanism. To distinguish between these two possible mechanisms more detailed investigation of the discourse-level processes is necessary. Within the eye-tracking context this would require a fine-grained analyses of the moments at which the lengthening of reading times and the regressions occur; other methods with high temporal resolution, such as ERP/EEG, may also be useful. Regardless, the current results do show, however, that context and background knowledge interact very early in the processing of incoming information and together constrain validation. Such conclusion is in line with spread-of-activation mechanisms posited in the discourse comprehension literature, such as the memory-based processing view (e.g., Cook et al., 1998b; Gerrig & McKoon, 1998; Myers & O'Brien, 1998; O'Brien & Myers, 1999; Rizzella & O'Brien, 2002) and cohort activation within the Landscape model view (van den Broek et al., 1999; van den Broek & Helder, 2017).

In this context, the observation in the current study that inaccuracy with world knowledge had stronger and longer effects than incongruency with context may reflect a structural property of the monitoring mechanisms – that validation always occurs first or primarily against the reader's background knowledge. It is also possible, however, that the observed dominance of world-knowledge is not an inherent property of the system but emerged due to other factors. For example, the dominance of one informational source over the other may depend on the strength of the reader's text-relevant general world knowledge (Cook, 2014) versus the strength of the contextual information (e.g., Cook & Guéraud, 2005; Myers et al., 2000; O'Brien & Albrecht, 1991). In the current study, knowledge inaccuracies tended to be stronger than the text incongruencies, as the former were outright errors whereas the latter merely were unlikely. To determine whether background knowledge structurally is

dominant, future studies could systematically vary the strength of background knowledge, similar to studies have varied the strength of the context (e.g., Creer et al., 2018; Walsh et al., 2018)

Early and late processes in validation

Theoretical models of validation assume distinct components to 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). Models such as the RI-Val model (Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b) focus on the passive, memory-based processes that are presumed to be involved in the initial detection of an inconsistency. Once detected, inconsistencies may trigger further processes, for example processes aimed at repairing the inconsistency (as described in the two-step model of validation; Isberner & Richter, 2014a; Richter, 2011; Richter et al., 2009; Schroeder et al., 2008). The models are not specific with respect to the relation between these components (e.g., does the detection component finish before possible repair processes, do the two components overlap, do detection processes interact with post-detection processes by triggering renewed detection processes?) but they generally agree that, as processing proceeds, the balance gradually shifts from detection to post-detection (repair) processes. Thus, although all eye movements may be influenced by both components of validation, early eye-tracking measures such as first-pass reading times are considered to reflect early processing (e.g., Clifton Jr et al., 2007; Keith Rayner & Liversedge, 2012) and therefore are relatively close to the detection processes. Conversely, later eye-tracking measures such as rereads and spill-over effects on subsequent sentences reflect later processing and are relatively more sensitive to reader-initiated (including possible repair) processes.

The current findings show that text-based and knowledge-based validation processes follow distinct trajectories in the very early stages of the processing of incoming information. Knowledge-based validation influences all early processes considered in this study, validation against earlier text also influences these processes but in qualitatively different ways depending on the presence or absence of knowledge violations. If the textual information is incongruent with the preceding text but fits the reader's background knowledge then the reader is likely to re-inspect the textual information. In contrast, if the textual information is incongruent with prior text and also violates the reader's background knowledge then the combined inconsistencies lead to longer reading time (over and above the already longer time due to the background knowledge inaccuracy), possibly reflecting more pervasive checking of textual input with background knowledge.

Interestingly, whereas initial text-based and knowledge-based validation processes show different processing patterns, later text-based and knowledge-based validation processes (e.g., regression path duration, re-reading probability, second-pass reading time, and several measures on the spill-over sentences) seem relatively similar. In so far as that later processing measures reflect repair processes, results suggest that repair processes for both types of inconsistencies involve a similar pallet of actions and sources. This may reflect that the final, adjusted mental representation of readers must fit with both contextual information and the existing knowledgebase. It is worth noting that the processing of knowledge inaccuracies required a more intensive, prolonged validation process (in line with van Moort et al, 2018), reflected in the presence of some spill-over effects (i.e., effects on first-pass reading of the spill-over sentence) for inconsistencies with background knowledge but not with text. Thus, knowledge inaccuracies in our study seemed more difficult to repair than textual incongruencies. It could be that in the case of knowledge inaccuracies the information has to be validated against a more elaborate network (i.e., the existing knowledgebase) than the episodic memory trace of the text representation and therefore it may take longer to activate relevant information. As mentioned above, this could also be caused by differences in strengths for the two types of inconsistencies.

In all, the results provide compelling evidence that the source of the incoherence influences processing from a very early stage. Both types of inconsistencies are detected early in processing with each triggering different processes. In comparison, in later processing the toolbox of (repair) processes for text incongruencies and knowledge inconsistencies seems rather similar.

The role of working memory in validation

The findings indicate that working memory modulates later processing (i.e., regressions, rereading): readers with a larger working-memory capacity made fewer regressions and were less likely to reread targets than those with smaller working-memory capacity. This suggests that readers adapt their later processing strategies depending on the resources they have available. Speculatively, readers with a larger working-memory capacity may have more relevant information available for processing, enabling them to avoid costly re-reading and reprocessing. In contrast, readers with a smaller working-memory capacity have less relevant information activated and, thus, may need to look back in the text to construct a coherent situation model.

No working-memory capacity effects were observed for the early processes. This suggests that early validation processes require few resources and is consistent

with the notion that such processes are relatively passive (e.g., RI-Val and two-step model of validation) whereas later validation processes are more resource demanding and reader-initiated.

In an earlier study using sentence-by-sentence self-paced reading, Van Moort et al. (2018) did observe an effect of working-memory capacity on reading times for knowledge inaccuracies but not text incongruencies, suggesting that knowledge-based validation is, in fact, resource demanding. Because the studies used the same materials but differed in presentation mode (sentence-by-sentence vs. texts presented in their entirety), it seems plausible that the constraints imposed by presentation mode may account for the different patterns of results. For example, during sentence-by-sentence presentation readers cannot look back to related information to resolve an inconsistency. Therefore, they may attempt to validate information for each sentence immediately and meticulously before proceeding in the text (Chung-Fat-Yim et al., 2017; Koornneef et al., 2019; Koornneef & Van Berkum, 2006) and, also, may need to rely more on their memory representation to conduct the validation (Gordon et al., 2006). As a result, sentence-by-sentence reading may elicit a greater effect of differences in working-memory capacity than reading of a text presented in its entirety. The potential effect of presentation mode on discourse-level comprehension processes is worth closer scrutiny as it would have important consequences for the interpretation of results from studies using sentence-by-sentence reading.

Broader implications and future directions

Successful comprehension requires readers to build a coherent, meaningful mental representation or situation model of a text. An essential aspect of building such mental representation is that readers routinely validate to what extent incoming information is consistent with what they already know. The current study shows that the processes involved in coherence monitoring depend on validation against both contextual information and background knowledge. Moreover, these sources exert their influence very early in the processing of new text information and they do so in distinct ways. The current conclusions are consistent with but also expand considerably current models of validation (e.g., RI-Val; Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b; The two-step model of validation; Isberner & Richter, 2014). They also are consistent with neuroimaging findings (Chapter 3 of this dissertation), revealing brain regions that seem mostly involved in either knowledge-based processing (e.g., dorsomedial Prefrontal Cortex) or text-based processing (right Inferior Frontal Gyrus) and regions that are affected by the two sources of information interactively (e.g., Precuneus and left Inferior Frontal Gyrus).

In addition to the text features that were the focus of this study, individual differences are likely to affect monitoring processes, especially in the later components of validation. We considered working memory, which indeed seemed to impose a capacity constraints on these later processes. Other individual difference factors that may differentially affect monitoring and repair processes are the standards of coherence that the reader applies or the toolbox of repair strategies he/she has available.

Comprehension monitoring occurs in the context of reading, as in this study and most validation research, but of course also in many other contexts of life. When encountering (fake) news, for example, one needs to validate whether the news is internally congruent and accurate with respect to world knowledge. Paradigms and models such as those discussed in this paper may provide a fruitful starting point for investigations of people's susceptibility to such (un)reliable information sources.

4