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

3

What you read vs what you know: Neural correlates of accessing context information and background knowledge in constructing a mental representation during reading

This chapter is based on:
van Moort, M. L., Jolles, D. D., Koornneef, A., & van den Broek, P. (2020). What you read vs what you know: Neural correlates of accessing context information and prior knowledge in constructing a mental representation during reading. *Journal of Experimental Psychology: General*, 149(11), 2084–2101.

Abstract

A core issue in psycholinguistic research is what the online processes are by which we combine language input and our background knowledge to construct the meaning of a message. We investigate this issue in the context of reading. To build a coherent and correct mental representation of a text readers monitor incoming information for consistency with the preceding text and with their background knowledge. Prior studies have not distinguished between text-based and knowledge-based monitoring, therefore it is unclear to what extent these two aspects of text comprehension proceed independently or interactively. We addressed this issue in a contradiction paradigm with coherent and incoherent versions of texts. We combined behavioral data with neuroimaging data to investigate shared and unique brain networks involved in text-based and knowledge-based monitoring, focusing on monitoring processes that affected long-term memory representations. Consistent with prior findings, behavioral results indicate that text and background knowledge each have a unique influence on processing. However, neuroimaging data suggests a more nuanced interpretation: Text-based and knowledge-based monitoring involve shared and unique brain regions, as well as regions that are sensitive to interactions between the two sources. It appears that the (d)mPFC and hippocampus -which are important for the influence of existing knowledge on encoding processes in non-reading contexts- are particularly involved in knowledge-based monitoring. In contrast, the right IFG is primarily involved in text-based monitoring, whereas left IFG and precuneus are implicated in integration processes. Furthermore, processes during reading affect recall of information (in)consistent with prior text or background knowledge.

Introduction

Constructing meaning (i.e., a mental representation that makes sense) from discourse is a fundamental human ability. To comprehend the world around us we build mental representations in which we integrate the current input (or recently acquired knowledge) with our existing knowledge base (stored in long term memory), for example when we read a book, watch a movie, or have a conversation. Building this representation is a dynamic process; the representation must be monitored and updated continuously as new information is encountered (e.g., Graesser et al., 1994; Kintsch & van Dijk, 1978; Trabasso et al., 1984; van den Broek et al., 1999). Moreover, it is crucial that the incoming information is validated to protect the mental representation against incongruencies or inaccuracies (e.g., Isberner & Richter, 2014a; Singer, 2013). The current study investigates these validation processes in the context of reading comprehension. Validating (written) materials against various sources of information is increasingly important, as they frequently contain inconsistencies, misinformation, or even fake news, especially today. Current theoretical frameworks presume a rudimentary cognitive architecture for validation processes, but they do not provide detailed information on when and how different sources of information, such as recently acquired knowledge (from the text) and readers' background knowledge (from memory), exert their influence. The current study examines the (neuro)cognitive architecture of the processes involved in text-based and knowledge-based validation.

Behavioral correlates of text-based and knowledge-based monitoring

Constructing meaning is a crucial ability in many major cognitive tasks, including learning, memory, perception, decision making, and language processing. The processes by which such meaning construction takes place have been investigated extensively in the context of language processing, in particular reading, as the interpretation of (written) language draws on many cognitive processes involved in meaning construction (Graesser et al., 1997). Although frameworks that describe how we construct meaning from language vary in approach and scope, they all attempt to describe the architecture and time course of processing: how and when the different language processes interact with each other. Some focus on the interplay between syntactic and extra-syntactic systems in processing individual sentences or very short discourse (e.g., Friederici, 2002; Hagoort, 2003), whereas others focus on how readers construct mental models when they process more extended discourse or texts (e.g., Kintsch, 1988; van den Broek et al., 1999; Zwaan et al., 1995).

A core aspect of constructing mental models from texts is the interplay between semantic and linguistic cues in the text (the text base) and relevant portions of world knowledge (which is integrated into the final mental representation) (e.g., Fodor, 1983; Kintsch, 1988; Millis & Just, 1994; van den Broek & Helder, 2017). Current theoretical frameworks of text comprehension provide elaborate descriptions of the cognitive processes involved, however they are underspecified on how and when various sources of information, such as contextual information or background knowledge, exert their influence on processing. For example, two prominent models of validation, the two-step model of validation and the RI-Val model, describe validation as the process where incoming information is evaluated for consistency with stored knowledge (Isberner & Richter, 2014a) or all information activated from long-term memory (Cook & O'Brien, 2014; O'Brien & Cook, 2016a, 2016b). This stored knowledge or activated information includes both information from the episodic representation of the text as well as general world knowledge. Although these models state that the two sources of information can impact validation processes, they do not describe the underlying architecture of text-based and knowledge-based validation processes. As a result, it is unclear whether validating against background knowledge and validating against prior text involve a common mechanism or (partially) separate mechanisms, and what happens when they are in conflict.

Interestingly, models that focus on processing individual sentences or very short discourse describe very similar issues (e.g., the interplay between syntax and semantics) but provide more detailed descriptions of the cognitive architecture of processing, and specifically, when and how various sources of information (e.g., syntax, semantics) influence processing (e.g., Ferreira & Clifton, 1986; Friederici, 2002; Hagoort, 2005; Hagoort et al., 2004; Hagoort & van Berkum, 2007; Jackendoff, 2007; van Berkum et al., 1999). To achieve such detailed descriptions neurobiological data has proven useful in developing, selecting, and constraining cognitive models, as assumptions on the cognitive architecture of a process also have implications for the hypothesized neural organization of that process (Hagoort, 2017). Analogous to models of sentence comprehension, models of text comprehension could benefit from neuroimaging research: Information on the neural architecture can be used to make more specific and more grounded claims on the cognitive architecture of text processing. More specifically, it can be used to make more specific claims on how different sources of information -such as contextual information and background knowledge- influence processing.

There is a considerable amount of behavioral data on the effect of contextual information and background knowledge on discourse processing (e.g., Albrecht & O'Brien, 1993; Nieuwland & Van Berkum, 2006; O'Brien & Albrecht, 1992; Rapp, 2008; Richter et al., 2009). The influence of these two sources of information on integration and updating processes is often examined using contradiction paradigms where participants read both coherent and incoherent versions of sentences or texts (e.g., Albrecht & O'Brien, 1993; Rapp, 2008). A robust finding in these paradigms is that readers slow down when they read sections that are inconsistent with their world knowledge, suggesting that they check the incoming information against background knowledge on the fly (O'Brien & Albrecht, 1992; Rapp, 2008). Background knowledge affects sentence and discourse processing at various levels of language processing (i.e., sentence and discourse level) and in different text genres (e.g., narrative and expository texts) (e.g., O'Brien & Albrecht, 1992; Rapp, 2008; Richter et al., 2009; Rodd et al., 2016; Wiley et al., 2018). Similarly, contextual information (i.e., provided within a text) can also affect processing and comprehension (e.g., Colbert-Getz & Cook, 2013; Myers et al., 2000). In fact, if contextual information is strong (e.g., recently encountered or more elaborate) it may even override the influence of world knowledge during reading ((e.g., Colbert-Getz & Cook, 2013; Myers et al., 2000; Nieuwland & Van Berkum, 2006; Rizzella & O'Brien, 2002). For example, Nieuwland and Van Berkum (2006) used a paradigm in which they presented readers with the sentence 'The peanut is in love' or 'The peanut is salted'. Although both are grammatically correct, the first sentence is more difficult to process than the second sentence because it violates background and lexico-semantic knowledge – i.e., peanuts can be salted but they are inanimate objects and therefore cannot be in love. However, when the same sentences were presented in a story about a peanut singing a song about his new girlfriend, readers would show the opposite pattern, because 'the peanut is in love' is appropriate giving the context, whereas 'the peanut is salted' is not.

Taken together, there is ample evidence that both contextual information and background knowledge influence the construction of a mental representation. However, studies on text-based monitoring do not always control for the influence of background knowledge to investigate the influence of prior text (Isberner & Richter, 2014a; van Moort et al., 2018). Instead, the influence of these two sources on processing and comprehension often are studied in tandem, for example in paradigms where the target (e.g., children are building a snowman) can only be incoherent with the preceding context (e.g., it was a hot sunny day) if readers had certain background knowledge (e.g., snow melts on a hot sunny day). As a result, it is unclear whether the mechanisms of validating against background knowledge are the same as those validating against prior text or whether these mechanisms are fundamentally different.

Addressing this issue, van Moort et al. (2018) developed a paradigm that explicitly contrasts validation against background knowledge and validation against prior text. Participants read expository texts about well-known historical topics containing a target sentence that was either true or false relative to the readers background knowledge and that was either supported or called into question in the preceding context. Reading time measures on targets provided a measure of readers' difficulty integrating statements into their discourse representation (e.g., Albrecht & O'Brien, 1993; Cook et al., 1998b; Rapp, 2008; Rapp et al., 2001). Results indicated that both prior text and background knowledge influenced readers' moment-by-moment processing on the target sentence, but only inconsistencies with background knowledge elicited a spill-over effect on the next sentence. This suggests that text-based monitoring and knowledge-based monitoring follow distinct time courses. Based on these results, Van Moort et al. (2018) speculated that text-based and knowledge-based monitoring may involve different cognitive systems. However, an alternative explanation suggests that there is a single system involved with these different types of validation, and the observed differences in reading times are the result of quantitative rather than qualitative differences in cognitive demands. Because the behavioral data is subject to multiple interpretations, neuroimaging data may help to better characterize the cognitive architecture of text-based and knowledge-based validation. More specifically, the main goal of the current study is to investigate whether we should distinguish between the influence of contextual information and background knowledge and, more specifically, whether we should assign separate roles for the two sources in the cognitive architecture of validation. By combining behavioral measures with neuroimaging data we aim to examine to what extent validation against background knowledge and validation against prior text call on the same underlying brain systems or (partly) different brain systems, as the neural organization of these processes can help differentiate between cognitive theories about common versus separate validation mechanisms (e.g., Frank & Badre, 2015; Hagoort, 2017). As background, we highlight some crucial findings in the neurocognitive literature below.

Division of labor in the coherence-monitoring network

Functional magnetic resonance imaging (fMRI) studies have begun to reveal a network of regions that contribute to the construction of coherent mental representations of texts (e.g., Egidi & Caramazza, 2013; Ferstl et al., 2008; Ferstl & Von Cramon, 2001; Mason & Just, 2006; Moss & Schunn, 2015; Virtue et al., 2006; Yarkoni et al., 2008). A meta-analysis of neuroimaging studies on text comprehension processes showed a network of regions that was more active for coherent compared to incoherent (or less coherent) language, including the left inferior frontal gyrus (IFG), medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC), precuneus, and several temporal lobe regions (Ferstl et al., 2008). These areas are thought to be involved in coherence building processes (Ferstl & Von Cramon, 2001, 2002; Kuperberg et al., 2006; Maguire et al., 1999; Mellet et al., 2002). However, as most studies in the meta-analysis compared processing coherent stories with processing sets of unrelated sentences it is difficult to determine to what extent regions in this network contribute to specific processes involved incoherence monitoring, such as detecting inconsistencies with prior text or background knowledge.

To study coherence monitoring in more detail, fMRI studies have employed variations of the contradiction paradigm (e.g., Ferstl et al., 2005; Hasson et al., 2007; Helder et al., 2017; Menenti et al., 2009). Two of these studies are particularly relevant for the current study. First, Helder, van den Broek, Karlsson, and Van Leijenhorst (2017) examined the neural correlates of coherence-break detection during reading and found a network of regions that were more active in response to incoherent than coherent target sentences, including the left IFG, precuneus, (dorso)medial prefrontal cortex (dmPFC), right supramarginal gyrus, and a number of temporal lobe regions. These findings suggest that the coherence-building network (Ferstl et al., 2008) becomes more active when inconsistencies are encountered. In addition, they found activation in subcortical clusters, including left hippocampus (HC), left amygdala, and bilateral caudate. Because the (in)coherence of targets could only be established in the context of readers' background knowledge the authors speculated that the left HC may have played a specific role in the reactivation of episodic memory traces of the text in combination with the retrieval of background knowledge. However, the paradigm they used did not allow distinguishing between these sources of information. Second, Menenti et al. (2008) investigated how context information provided within a text influences the processing of (erroneous or implausible) world-knowledge information. Their results suggest that the left IFG is particularly sensitive to the consistency of information with background knowledge (showing increased activation for false information), whereas the right IFG also takes into account the ongoing discourse (showing a reduced effect of false information if the context provides an explanation) (Menenti et al., 2008). These findings suggest that both world knowledge and discourse context affect the integration of world knowledge into the mental representation, and do so by recruiting partly different sets of brain areas.

As in the behavioral research, most neuro-imaging studies do not make a clear distinction between text-based and knowledge-based monitoring. The study by Menenti et al. (2008) was the only study that included contextual and world knowledge manipulations in a single design, but their main goal was to examine whether the context can override world knowledge, thereby disregarding text-based monitoring as an independent process. As a result, it is difficult to disentangle potentially separate monitoring networks. However, these studies do suggest that the regions involved, i.e., (d)mPFC, precuneus, left and right IFG, display a division of labor with respect to building, monitoring, and updating the situation model (e.g., Ferstl et al., 2008; Hagoort et al., 2004; Helder et al., 2017; Menenti et al., 2009).

The dmPFC and the precuneus are both key nodes of a 'default-mode network' that is involved in, among other processes, building mental models (e.g., Andrews-Hanna et al., 2010). During reading, these regions are suggested to be important for monitoring and inferencing in complex or ambiguous situations as well as in situation-model building (Ferstl et al., 2008). Their functional division is still under debate, but it seems that the dmPFC may be particularly important for coherence monitoring (e.g., Ferstl et al., 2005; Hasson et al., 2007; Helder et al., 2017), whereas the precuneus/PCC may play a crucial role in building and updating mental representations (e.g., Speer et al., 2009).

The left IFG is a key region in a left lateralized perisylvian language network that is associated with basic language processing on the word and sentence level, as well as with more complex processes such as coherence building (Ferstl et al., 2008). As described above, it is sensitive to sentences that are incoherent with the text given that readers have certain background knowledge (e.g., Helder et al., 2017) and is involved in recruiting world knowledge during comprehension (e.g., Hagoort et al., 2004; Menenti et al., 2009). Hagoort et al. (2004) showed that sentences participants regarded as false elicited an activity increase in the left IFG compared to true sentences. The left IFG is sensitive to the consistency of information with background knowledge, even if the preceding discourse overrides this knowledge (Menenti et al., 2008). So, although the left IFG seems to play a role in both text-based and knowledge-based monitoring, it may be more dedicated to world-knowledge validation processes. Similar claims have been made for the right IFG, although this region is thought to play a larger role in integrating previously stored knowledge with discourse information (Menenti et al., 2008).

In summary, both prior text and background knowledge influence the neurocognitive processes that take place during reading comprehension. These two informational sources are processed in overlapping neural networks, consisting of several core brain regions (dmPFC, precuneus/PCC and left and right IFG). It is not year clear whether these core regions are uniquely involved in either text-based or knowledge-based processing, or whether they contribute to both types of validation. Therefore, the second goal of the current study is to provide a more detailed picture of the division of labor between these regions during text-based and knowledge-based monitoring.

Updating the memory representation

When new information is encountered during reading, the long-term memory representation can be updated or revised to take this information into account (Kendeou, 2014). However, it is not yet clear how reading an inconsistency affects the long(er) term memory representation. Therefore, the third goal of the current study is to investigate the influence of text-based and knowledge-based monitoring on the long-term memory representation. With respect to memory performance, reading inconsistent information could affect the memory representation in different ways: It could be that information that fits with pre-existing memory representations (i.e., the readers' background knowledge or the mental representation of the text) is remembered better (e.g., Anderson et al., 1983). Alternatively, it could be that inconsistent information is remembered better, as it is relatively 'new' information with respect to the mental representation or the readers' background knowledge and novelty is suggested to improve retention of that information (e.g., Kormi-Nouri et al., 2005). These possible memory performance differences may be reflected in differences in neural activation during encoding. This is illustrated by a study by Hasson et al. (2007), who showed that activation in the precuneus, dmPFC and right superior temporal gyrus predicted memory for spoken narratives. This study suggested that at least part of the coherence-monitoring network is involved in both detecting (and repairing) inconsistencies and in encoding them in memory. Thus, differences in memory encoding between inconsistent and consistent information may affect the neural activation during the task, therefore it is important to take memory performance into account.

The current study

As current theoretical frameworks of text comprehension are underspecified on when and how contextual information and background knowledge influence validation processes, it is unclear whether we should distinguish between the influence of contextual information and background knowledge and, more specifically, whether we should assign separate roles for the two sources in the cognitive architecture of validation. Moreover, models of text comprehension are not grounded in the neural architecture of the brain, as neuroimaging research investigating coherence monitoring is still in its early stages. To investigate the specific roles of contextual information and background knowledge in the cognitive architecture of validation processes and explore the neural underpinnings of these processes, the current study combined behavioral data with neuroimaging data in a contradiction paradigm with coherent and incoherent versions of text. To investigate to what extent text-based and knowledge-based monitoring call on the same or (partly) different brain regions we investigated specific regions of interest that are known to play a key role in coherence monitoring (i.e., dmPFC, precuneus, left and right IFG) to examine a possible division of labor between them. Moreover, we investigated whether other relevant regions for text-based and knowledge-based monitoring could be identified and compared the results to studies that (tentatively) proposed a more extended coherence monitoring network (e.g., Helder et al., 2017). Finally, to allow for a more grounded interpretation of the neural correlates of online text processing and later recall, we investigated whether the processes during initial reading affect the recall of information (in)consistent with prior text or background knowledge. In addition, as these possible differences in memory encoding may affect the neural activation during the task, we took memory performance into account in our fMRI analyses by focusing specifically on successfully retrieved targets. Therefore, we only included trials in our analyses if the targets were remembered correctly as we can be confident that in these trials the text representation was updated successfully.

We employed a self-paced sentence-by-sentence reading paradigm, with recording of reading times and neuroimaging data during reading. Reading times provide a behavioral measure of readers' difficulty integrating statements into a mental representation as texts unfold (e.g., Albrecht & O'Brien, 1993; Cook et al., 1998b). Participants read expository texts in which half of the texts included a target sentence that contained information that could be either true or false with the readers background knowledge. At the same time, the context sentences prior to the target provided contextual information that made it more or less likely that the event described in the target sentence occurred, while remaining historically accurate (see Table 1 for a sample text). In line with previous behavioral studies (e.g., Albrecht & O'Brien, 1993; Cook et al., 1998b; Rapp, 2008) we expected longer reading times for the targets that were inconsistent with either text or background knowledge, which would suggest that readers indeed check the incoming information against background knowledge and the preceding text. To examine the respective influences of information from the text and from background knowledge on the mental representation and participants' memory of the texts, we employed a surprise memory task the next day. The aim of this memory task was twofold: First, recall data would give us insight in the effects of processing inconsistencies on the memory representation. Second, it allowed us to elaborate on the neurocognitive processes involved in updating the mental representation of a text during reading by specifically considering the fMRI results for correctly remembered items.

Method

Participants

Thirty-one, right-handed, native speakers of Dutch (11 men, 20 women) aged 19-28 years (*M* = 23, *SD* = 3) participated for monetary compensation. All participants had normal eyesight and none reported having neurological and/or psychiatric disorders or using psychotropic medication. Participants provided written informed consent. Procedures were approved by the internal review board at the Leiden University Medical Centre and conducted in accordance with the Declaration of Helsinki. Anatomical scans were cleared by a radiologist.

Materials

The 40 texts were a subset of the materials used in Chapter 2 of this dissertation (based on Rapp, 2008). The texts were normed to ensure that the presented facts were common knowledge in our sample. The presented facts were familiar to at least 80% of the participants in the norming study (see Chapter 2 of this dissertation for a more detailed description of the norming study). Furthermore, only on 1.5% of the items participants in the current study indicated that they did not have any or very little knowledge of the topic on a questionnaire assessing their background knowledge they filled in afterwards. The texts are about well-known historical topics and each text contains a target sentence that is either true or false (with the readers' background knowledge); at the same time the preceding text could either support (i.e., unambiguous context) or call into question (i.e., suspenseful context) the information in the targets. Four different versions of each text were constructed by orthogonally varying the target sentence itself (true vs. false) and the context prior to the target (congruent vs. incongruent with target). 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 3.1). It is important to note that bias-false contexts 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 3.1). 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” 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.”). Sentences 9-10 concluded the text and did not elaborate on the fact potentially called into question in the target. 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 target (true or false) and text context (congruent or incongruent with target) on each list. Each participant was assigned to one list and, hence, read one version of each text.

Table 3.1. 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.

Apparatus

Reading task

Participants read the 40 texts in the scanner in two blocks. Texts were presented sentence-by-sentence, while reading times were recorded. The presentation rate was self-paced. Participants were instructed to read for comprehension (“Please read the texts carefully, it is important that you understand the texts”) and to advance to the next sentence by pressing a button with their left index finger. Sentences remained on screen for a maximum of 10 seconds. A fixation cross was presented between texts (variable duration 7000 – 14800 ms). One second before the next trial would start the cross would turn red to alert participants to the start of the trial. A short practice block preceded the experimental blocks.

Recognition Memory Task

The recognition memory task consisted of 160 sentences: 40 target, 40 context, 40 neutral, and 40 distractor sentences. Participants were presented with sentences that either matched or did not match sentences they had encountered in the reading task (e.g., when they were presented with ‘the statue of liberty was delivered to the US’ during the reading task they could be presented with either ‘the statue of liberty was delivered to the US’ or ‘the statue of liberty was not delivered to the US’). For each sentence participants indicated whether they had read the sentence during the reading task in the scanner or not and rated confidence in their answer on a scale ranging from (1) very uncertain to (7) very certain. 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 (i.e., 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 have 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 were not presented at all in the reading task.

Procedure

Participants were tested individually in two sessions. In the first session they completed the reading task in the MRI scanner (total duration ca. 75 min). The second session took place about 24 hours after the first. In this session they completed the memory task and a questionnaire to assess whether they had the required background knowledge or not by indicating for each topic how much they knew about it prior to reading the text on a scale from 1 (nothing) to 7 (a lot).

Behavioral data analysis

To investigate the effects of the manipulations on the reading process we conducted mixed-effects linear regression analyses on the log-transformed reading times for target and spill-over sentences (sentence following the target). For 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 with preceding context) and the random factors subjects and items. Based on the results of this analysis we selected the most parsimonious model and included the maximal participant and item random-effect structure that resulted in a converging model (Barr et al., 2013). To investigate the effects of the manipulations on participants' recall of the texts we conducted mixed-effects logistic regression analyses on memory and certainty scores using the same approach. We will only report final models for each analysis. 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). We report both likelihood tests and tests of the fixed estimates for all models. Unless mentioned otherwise, we only discuss effects that were significant in both likelihood tests ($p < 0.05$) and fixed-estimates tests ($t > 1.96$).

fMRI data acquisition

fMRI data were acquired using a standard whole-head coil on a 3-Tesla Philips Achieva MRI scanner. Foam inserts were used to minimize head movement. Prior to the functional runs, a high resolution 3DT1-weighed anatomical scan was obtained for registration purposes (repetition time (TR) = 9.7; echo time (TE) = 4.60, flip angle = 8%, 140 axial slices, field of view (FOV) = 224 mm × 168 mm × 177.333 mm, 0.275mm slice gap and voxel size = 0.875 mm × 0.875 mm × 1.2 mm).

For fMRI, T2*-weighed whole-brain Echo-Planar Images were acquired in two runs with the following parameters: TR = 2.2 s, TE = 30 ms, flip angle = 80°, 38 axial slices, FOV = 220 × 115 x 220 mm, 2.75 mm isotropic voxels, 0.275 mm slice gap, voxel size = 2.75 mm × 2.75 mm × 2.75 mm and max 800 volumes per run. Five dummy scans preceded each run to allow for equilibration of T1 saturation effects. Because the task was self-paced the number of volumes per block varied. When the participants finished the task the T2 scan was stopped. Stimuli were projected using E-prime software (version 2.0.10.147, Psychology Software Tools) onto a screen at the head of the scanner bore which participants viewed through a mirror attached to the head-coil.

fMRI preprocessing and data analyses

Data pre-processing was performed using FSL (version 5.0.9) and consisted of motion correction using MCFLIRT (Jenkinson et al., 2002), non-brain removal using BET (Smith, 2002), 5 mm Gaussian kernel FWHM spatial smoothing, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 100 s). Functional scans were registered to the T1-weighted images, and then to the 2 mm MNI-152 standard space template.

Data analysis was performed using FEAT (version 6.00), part of FSL. A set of whole-brain analyses was conducted to identify regions involved in text-based and knowledge-based monitoring. We focused specifically on trials that were remembered correctly in the memory task because for these trials we can be confident that the text representation was updated successfully[3]. Trials that were remembered incorrectly in the memory task were included in the model but excluded from the contrasts of interest. For each participant in each run, ten EVs with their temporal derivatives were included in the general linear model, representing the presentation of (1) sentences 1-7, (2) congruent true target, correctly remembered, (3) congruent true target, incorrectly remembered, (4) congruent false target, correctly remembered, (5) congruent false target, incorrectly remembered, (6) incongruent true target, correctly remembered, (7) incongruent true target, incorrectly remembered, (8) incongruent false target, correctly remembered, (9) incongruent false target, incorrectly remembered and (10) sentences 9-10. Onset of the EVs was determined using custom-written scripts in R Studio (version 0.99.903, RStudio, Inc.), based on each participant's button presses. EVs were convolved with a double gamma hemodynamic response function.

---

[3] Analyses including all trials can be found in the supplementary material

At the single subject level, we created the following contrasts: (1) incongruent > congruent, (2) congruent > incongruent, (3) false > true, (4) true > false, (5) positive interaction, i.e., voxels where false > true is larger for congruent than for incongruent texts, (6) negative interaction, i.e., voxels where false > true is larger for incongruent than for congruent texts. We did not create contrasts for incorrectly remembered trials because some of the conditions did not have enough incorrect trials. A direct comparison between correct and false could not be performed for the same reason. Contrasts were combined across the runs on a subject-by-subject basis using fixed-effects analyses (Woolrich et al., 2004), and submitted to whole-brain mixed-effect group analyses. Resulting whole-brain statistical maps were corrected for multiple comparisons using cluster-based correction ($p < .0.05$, initial cluster-forming threshold $Z > 2.3$). All local maxima are reported as MNI coordinates. Anatomical location was determined using the Harvard-Oxford Cortical and Subcortical structural atlas for FSL.

ROI analyses were performed using Featquery and SPSS version 23 (IBM Corp., Armonk, NY, USA), focusing on dmPFC, IFG, and precuneus. We created 8 mm spherical ROIs centered at local maxima described by Helder et al. (2017): dmPFC [-9, 48, 21], precuneus [-12, -45, 42], left IFG [-42, 24, -12] and its right homologue [42, 24, -12]. Repeated-measures ANOVAs were conducted to examine the effects of text (congruent, incongruent) and background knowledge (true, false) on the mean activation in each ROI on correctly remembered targets.

Results

Reading task

To examine the effects of the manipulations we conducted mixed-effects linear regression analyses on the log transformed reading times on target and spill-over sentence (see Table 3.2 for descriptives). The Wald chi-square test revealed main effects of background knowledge ($\chi2$ (1) = 11.34, $p < 0.001$) and of text ($\chi2$ (1) = 8.05, $p = 0.005$): reading times were longer for false than for true targets ($\beta = 0.09$, $SE = 0.026$, $t = 3.37$) and reading times were longer for incongruent than for congruent targets ($\beta = 0.05$, $SE = 0.018$, $t = 2.84$). On the spill-over sentence the Wald chi-square tests revealed a marginal effect of background knowledge ($\chi2$ (1) = 3.39, $p = 0.066$), and no effect of text.

Table 3.2. Mean reading times and standard deviations (in ms) at the regions of interest (target and spill-over sentence) for the experimental manipulations of background knowledge (target true or false) and text (target congruent or incongruent with preceding context).

Text	**Background knowledge**	**Target**		**Spill-over**	
		M	***SD***	***M***	***SD***
Congruent	True	2715	1051	3370	1349
	False	2938	1118	3520	1348
Incongruent	True	2810	970	3343	1309
	False	3141	1301	3433	1236

Memory task

Participants were proficient in distinguishing whether they had read a sentence or not. Overall, they scored on average 77% correct and they were 75% certain of their answer. On sentences originating from the task (target, context, and neutral) they scored on average 73% correct and were 73% certain of their answer. On distractor sentences they scored on average 91% correct and were 72% certain of their answers. This shows they had read the texts attentively. To investigate the effect of the manipulations on memory of the texts we conducted mixed-effects linear logistic regression analyses on accuracy and certainty scores for target sentences. The Wald chi-square test revealed main effects of background knowledge ($\chi^2(1) = 8.327$, $p = 0.003$) and of text ($\chi^2(1) = 7.354$, $p = 0.006$) on the accuracy scores, indicating that false targets were remembered less well than true targets ($\beta = -0.86$, $SE = 0.298$, $z = -2.89$) and that targets that were incongruent with the preceding text were also remembered less well than congruent targets ($\beta = -0.58$, $SE = 0.215$, $z = -2.71$). We did not find any effects of text or background knowledge on the certainty scores.

Region-of-interest analyses

For the mean activation in the dmPFC we found a main effect of background knowledge ($F(1) = 8.85$, $p = 0.006$) but not of text ($F(1) = 0.48$, $p = 0.492$): false targets elicited more activation than true targets (Fig. 3.1). We observed a trend in the interaction ($F(1) = 3.95$, $p = 0.056$). For the mean activation in the precuneus we found a text x background knowledge interaction ($F(1) = 4.28$, $p = 0.047$), but no main effects of text ($F(1) = 0.09$, $p = 0.769$) or background knowledge ($F(1) = 0.21$, $p = 0.649$)

(Fig. 3.1). For the mean activation in the left IFG we found a text x background knowledge interaction (*F*(1) = 10.86, *p* = 0.003), but no main effects of text (*F*(1) = 0.42, *p* = 0.521) or background knowledge (*F*(1) = 1.12, *p* = 0.298) (Fig. 3.1). For the mean activation in the right IFG we found a main effect of text (*F*(1) = 5.24, *p* = 0.029) but not of background knowledge (*F*(1) = 3.29, *p* = 0.080): congruent targets elicited more activation than incongruent targets. We did not find a significant interaction (*F*(1) = 0.40, *p* = 0.842) (Fig. 3.1).

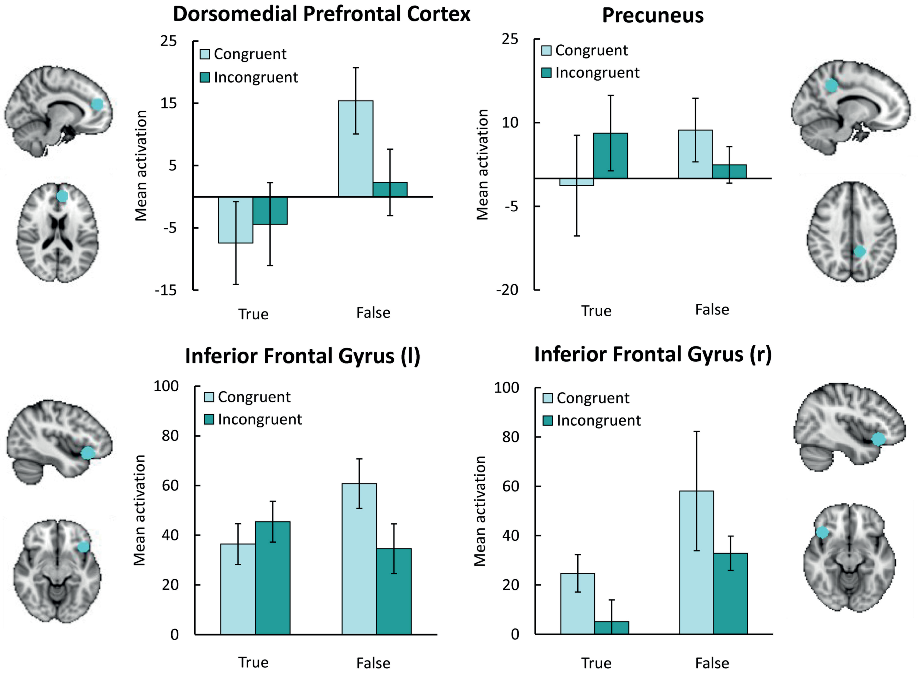

Figure 3.1. ROI mean activation in response to true or false targets and congruent or incongruent targets, for 8mm spherical ROIs centered at MNI coordinates [-9, 48, 21] (dmPFC), [-12, -45, 42] (precuneus), [-42, 24, -12] (left IFG), and [42, 24, -12]

Whole-brain analyses of text-based and knowledge-based monitoring

We conducted a set of whole brain analyses to explore the involvement of other regions in text-based and knowledge-based monitoring. To examine the neural correlates of knowledge-based monitoring, we contrasted activation on trials in which participants read true targets with activation on trials in which they read false targets and vice versa. The whole-brain contrast of true versus false resulted in clusters of activation in the left and right lateral occipital cortex extending into the fusiform gyrus, and the left HC (Table 3.3; Fig. 3.2). The reverse contrast yielded a cluster of activation in the dmPFC[4], left angular gyrus (AG), and bilateral lateral PFC. To examine the neural correlates of text-based monitoring, we contrasted activation on trials in which participants read targets that were incongruent with the context with activation on trials in which they read congruent targets and vice versa. The whole-brain contrast of congruent versus incongruent targets resulted in clusters of activation in bilateral inferior temporal occipital cortex, bilateral superior parietal lobule, precuneus[5], and supplementary motor cortex (Table 3.3; Fig. 3.2). The reverse contrast did not yield any significant clusters. The text by background knowledge interaction resulted in clusters of activation in the bilateral middle frontal gyrus (MFG), left IFG[6], and right AG (Table 3.3; Fig. 3.2; Fig 3.3). These regions showed more activation for incongruent than congruent targets if targets reflected true world knowledge information, but more activation for congruent than incongruent targets if targets reflected false world knowledge information. The reverse contrast did not yield any significant clusters.

---

4 The same portion of the dmPFC was activated in both whole-brain and ROI analyses.

5 A different portion of the precuneus was activated in the whole-brain contrast congruent > incongruent than in the ROI analyses.

[6] In the whole-brain analyses a more dorsal portion of the left IFG was activated than in the ROI analyses.

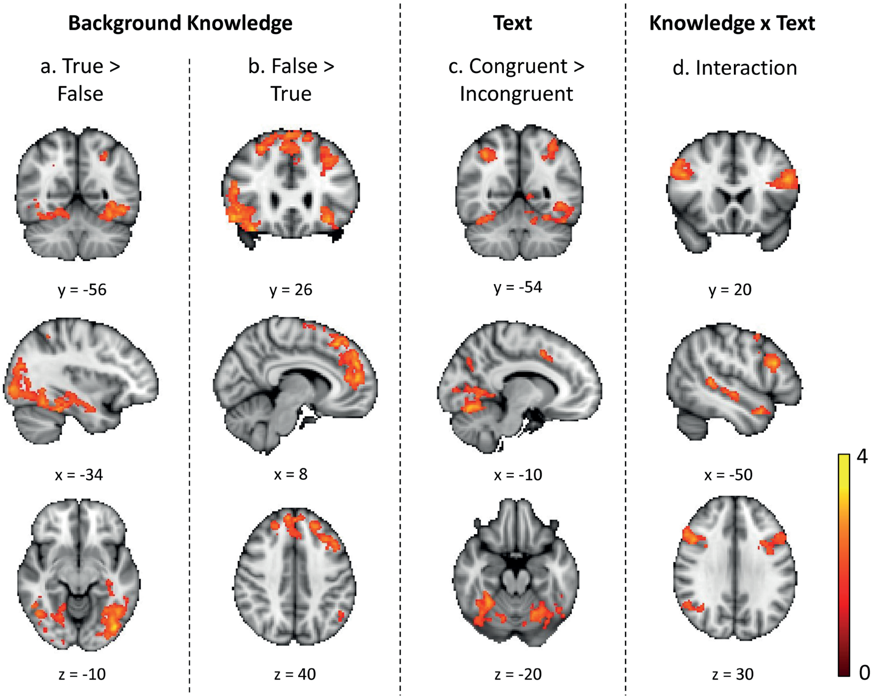

Figure 3.2. Whole-brain statistics maps for (a) the contrast true > false targets, (b) the contrast false > true targets, (c) congruent > incongruent targets, and (d) interaction text x background knowledge across all participants (thresholded at z =2.3 and p < 0.05). The left side of the brain is plotted on the right side of the image.

Table 3.3. Whole-brain group activations for the correctly remembered targets in response to true/false and congruent/incongruent targets.

Anatomical region	L/R	Z max	MNI coordinates X	Y	Z	voxels	*p*
a. Results for contrast true > false							
40% Inferior Lateral Occipital Cortex	L	3.85	-38	-72	4	3317	$p < 0.001$
57% Inferior Lateral Occipital Cortex	L	3.83	-42	-76	-8		
35% Superior Lateral Occipital Cortex	L	3.8	-22	-80	24		
48% Superior Lateral Occipital Cortex	L	3.63	-24	-82	28		
55% Posterior Temporal Fusiform Cortex	L	3.48	-32	-40	-16		
28% Occipital Fusiform Gyrus	L	3.42	-38	-70	-10		
39% Inferior Lateral Occipital Cortex	R	3.64	36	-84	4	1723	$p < 0.001$
46% Inferior Lateral Occipital Cortex	R	3.4	40	-82	-2		
62% Temporal Occipital Fusiform Cortex	R	3.33	30	-50	-14		
16% Temperooccipital Inferior Temporal Gyrus	R	3.28	42	-58	-8		
31% Inferior Lateral Occipital Cortex	R	3.27	46	-62	-8		
39% Occipital Fusiform Gyrus	R	3.1	22	-64	-12		
39% Superior Lateral Occipital Cortex	R	3.83	26	-64	34	535	0.008
60% Superior Lateral Occipital Cortex	R	3.45	32	-72	30		
58% Superior Lateral Occipital Cortex	R	2.96	34	-78	26		
58% Superior Lateral Occipital Cortex	R	2.88	24	-74	52		
73% Superior Lateral Occipital Cortex	R	2.82	30	-74	44		
31% Superior Lateral Occipital Cortex	R	2.81	20	-70	46		
b. Results for contrast false > true							
36% Superior Frontal Gyrus	R	3.86	2	36	54	3515	$p < 0.001$
17% Superior Frontal Gyrus	L	3.78	0	30	58		
62% Paracingulate Gyrus	R	3.74	2	44	30		
59% Paracingulate Gyrus	R	3.7	6	52	16		
27% Superior Frontal Gyrus	R	3.64	8	50	30		
66% Paracingulate Gyrus	R	3.63	6	52	12		
25% Inferior Frontal Gyrus, pars triangularis[7]	R	4.3	50	22	12	1667	$p < 0.001$
65% Frontal Orbital Cortex	R	4.18	28	20	-14		
57% Frontal Orbital Cortex	R	4.09	44	30	-12		
76% Frontal Orbital Cortex	R	4.08	34	24	-20		
36% Frontal Orbital Cortex	R	3.58	46	26	-4		
77% Frontal Pole	R	3.57	48	40	-14		

---

[7] Note that there is no overlap in activation in the right IFG between the whole-brain and the ROI results.

40% Posterior Supramarginal Gyrus	L	3.69	-58	-52	26	467	0.018
49% Angular Gyrus	L	3.32	-56	-58	30		
39% Posterior Supramarginal Gyrus	L	3.27	-64	-42	24		
46% Posterior Supramarginal Gyrus	L	3.2	-58	-50	34		
53% Superior Lateral Occipital Cortex	L	3.16	-56	-62	30		
45% Posterior Supramarginal Gyrus	L	3.07	-64	-44	28		
62% Middle Frontal Gyrus	L	3.26	-42	20	42	444	0.024
40% Middle Frontal Gyrus	L	3.26	-34	24	38		
35% Frontal Pole	L	3.22	-22	38	40		
48% Middle Frontal Gyrus	L	3.21	-44	20	38		
65% Middle Frontal Gyrus	L	3.15	-46	16	42		
30% Superior Frontal Gyrus	L	2.96	-26	28	48		
53% Frontal Orbital Cortex	L	3.53	-28	22	-14	414	0.034
36% Frontal Orbital Cortex	L	3.46	-32	26	-12		
37% Frontal Orbital Cortex	L	3.29	-28	12	-16		
c. Results for contrast congruent > incongruent							
22% Lingual Gyrus	L	3.57	-10	-72	-14	1538	$p < 0.001$
10% Lingual Gyrus	L	3.38	-18	-66	-16		
100% Cerebellum	L	3.3	-20	-60	-20		
42% Inferior Lateral Occipital Cortex	L	3.29	-46	-64	-4		
2% Temporal Occipital Fusiform Cortex	L	3.24	-26	-60	-22		
36% Temperooccipital Inferior Temporal Gyrus	L	3.21	-42	-54	-8		
60% Temporal Occipital Fusiform Cortex	R	3.42	30	-48	-20	865	$p < 0.001$
60% Temporal Occipital Fusiform Cortex,	R	3.42	32	-42	-22		
71% Inferior Lateral Occipital Cortex	R	3.27	44	-82	-10		
72% Temporal Occipital Fusiform Cortex	R	3.24	40	-50	-22		
61% Inferior Lateral Occipital Cortex	R	3.14	50	-68	6		
66% Inferior Lateral Occipital Cortex	R	3.09	54	-64	8		
65% Supplementary Motor Cortex	R	3.43	2	0	54	756	0.009
58% Paracingulate Gyrus	R	3.29	2	10	48		
34% Anterior Cingulate Gyrus	R	3.25	10	10	40		
66% Paracingulate Gyrus	R	3.23	6	16	46		
22% Paracingulate Gyrus	L	3.2	-10	8	44		
14% Paracingulate Gyrus	R	2.97	14	18	34		
50% Superior Parietal Lobule	L	3.16	-34	-52	52	586	0.005
40% Precuneous Cortex[8]	L	3.07	-14	-70	30		

---

[8] Note that the whole-brain and the ROI analyses showed activation in different portions of the precuneus.

13% Superior Parietal Lobule	L	3.04	-24	-46	42		
62% Superior Lateral Occipital Cortex	L	2.98	-26	-66	46		
55% Superior Lateral Occipital Cortex	L	2.95	-24	-68	42		
35% Superior Parietal Lobule	L	2.93	-34	-58	60		
26% Angular Gyrus	R	3.38	36	-54	44	449	0.025
54% Superior Lateral Occipital Cortex	R	3.29	28	-64	44		
41% Superior Parietal Lobule	R	2.92	36	-44	44		
33% Precuneous Cortex	R	2.88	16	-64	42		
19% Posterior Supramarginal Gyrus	R	2.87	38	-36	40		

d. Results for contrast incongruent > congruent

No significant clusters

e. Results for contrast interaction

24% Inferior Frontal Gyrus, pars opercularis	L	3.61	-50	22	24	966	*p* < 0.001
41% Middle Frontal Gyrus	L	3.38	-40	12	42		
62% Middle Frontal Gyrus	L	3.32	-42	12	52		
25% Inferior Frontal Gyrus, pars triangularis	L	3.31	-58	24	16		
8% Middle Frontal Gyrus	L	3.28	-30	10	36		
17% Inferior Frontal Gyrus, pars opercularis	L	3.17	-40	22	20		
39% Posterior Superior Temporal Gyrus	L	3.55	-60	-40	4	687	0.001
36% Posterior Superior Temporal Gyrus	L	3.42	-50	-40	6		
73% Temporal Pole	L	3.35	-50	8	-22		
36% Posterior Superior Temporal Gyrus	L	3.17	-52	-14	-8		
53% Posterior Middle Temporal Gyrus	L	3.02	-60	-38	-4		
41% Posterior Superior Temporal Gyrus	L	2.99	-50	-26	-4		
33% Middle Frontal Gyrus	R	3.62	46	18	28	472	0.013
44% Middle Frontal Gyrus	R	3.48	54	26	32		
43% Middle Frontal Gyrus	R	3.39	54	22	32		
61% Middle Frontal Gyrus	R	3.19	50	20	40		
55% Middle Frontal Gyrus	R	3.1	44	16	40		
13% Precentral Gyrus	R	3.04	58	16	32		
62% Angular Gyrus	R	3.24	52	-56	26	451	0.017
47% Angular Gyrus	R	3.15	56	-58	34		
34% Superior Lateral Occipital Cortex	R	3.13	42	-62	26		
46% Superior Lateral Occipital Cortex	R	3.04	50	-60	26		
31% Superior Lateral Occipital Cortex	R	2.96	60	-62	22		
50% Angular Gyrus	R	2.88	46	-50	20		

f. Results for contrast negative interaction

No significant clusters

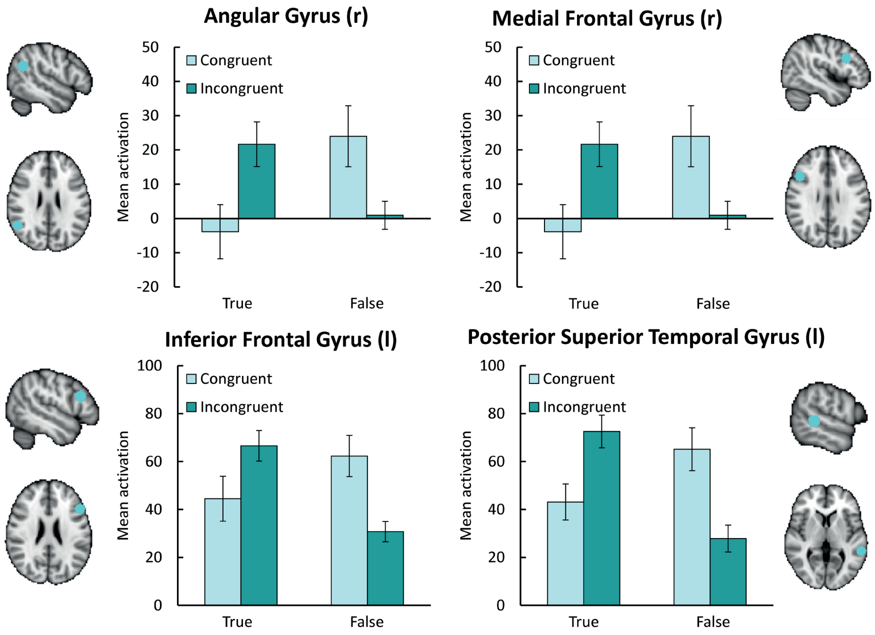

Figure 3.3. Activation diagrams for the four significant clusters in the whole-brain positive interaction. ROI mean activation in response to true/false and congruent/incongruent targets for peaks of the four significant clusters of activation in the whole-brain positive interaction contrast centered at MNI coordinates [52, -56, 26] (right AG), [46, 18, 28] (right MFG), [-50, 22, 24] (left IFG) and [-60, -40, 4] (left PSTG). The left side of the brain is plotted on the right side of the image.

Discussion

Although current theoretical frameworks of text comprehension are clear that contextual information and background knowledge influence processing, they are underspecified on *how* and *when* these sources influence processing. Therefore, the current study aimed to explicate these models by investigating whether we can distinguish between the influence of contextual information and background knowledge in coherence monitoring and whether we should assign separate roles for the two sources in the cognitive architecture of validation. Furthermore, we examined the neural correlates of text-based and knowledge-based monitoring to ground models of validation in the neural architecture of the brain and broaden our understanding of the neural underpinnings of coherence monitoring. Finally, we examined how reading inconsistent information affects the long(er) term memory representation to shed light on how validation processes during reading affect long-term memory.

In line with our prior work, behavioral results showed inconsistency effects for both text and background knowledge on targets, in the absence of an interaction between text and knowledge effects (van Moort et al., 2018). Furthermore, we found that validation processes also affected the long(er) term memory representation, i.e., mismatching information -either with the text or with background knowledge- is remembered less well than matching information. To examine whether the behavioral effects were driven by a singular or two separate validation systems, we studied the underlying brain networks involved in text- and knowledge-based processing. The strongest evidence in favor of a multiple systems account would be a double dissociation in terms of the brain regions involved in both types of validation. Our results provide some evidence of such a dissociation, but indicate that there are interactions between the systems as well. More specifically, our ROI data revealed a main effect of knowledge (but not of text) in the dmPFC, and a main effect of text (but not of knowledge) in the right IFG, and an interaction between knowledge and text in the precuneus and left IFG. These results were largely confirmed by whole-brain analyses. We found a large network of regions involved only in knowledge-based monitoring (including the mPFC and right HC, Table 3.3; Fig. 3.2), a somewhat smaller network of regions that were involved only in text-based monitoring (including the ACC, Table 3.3; Fig. 3.2), and a number of regions that showed an interaction (e.g., left IFG, right AG and bilateral MFG, Table 3.3; Fig. 3.2). Finally, In the whole brain data, there were also a number of regions where text-based and knowledge-based monitoring processes had similar effects, showing two main effects but no interaction (e.g., inferior temporal occipital cortex and superior parietal lobule).

Cognitive architecture of validation

Our first aim was to explicate the specific roles of contextual information and background knowledge in the cognitive architecture of validation and investigate whether we should assign separate roles for these two sources. Van Moort et al. (2018) suggested that prior text and background knowledge each have a unique influence on processing and that knowledge-based inconsistencies have a more pronounced effect on processing. Consistent with this notion, our behavioral results showed inconsistency effects for both text and background knowledge on targets, in the absence of an interaction between text and knowledge effects. However, in contrast to van Moort et al. (2018) our behavioral results did not show a more pronounced effect for knowledge-based inconsistencies, as we did not find a spillover effect of background knowledge.

To further differentiate between text-based and knowledge-based validation, we performed neuroimaging analyses. The results of these analyses are largely in line with the hypothesized multiple-systems account, but they paint a more nuanced picture. In line with the multiple-systems account, we found a number of regions that showed a main effect of either text or background knowledge, suggesting that readers process text-based and knowledge-based information separately, at least to some extent (e.g., in the right IFG, dmPFC, and HC). Furthermore, whole-brain results indicated that knowledge-based monitoring recruits a larger network of brain regions than text-based monitoring. Thus, whereas the behavioral results did not show a more pronounced effect for knowledge-based inconsistencies, the neuroimaging results did. There are a couple of possible explanations for this pattern of results. First, a process-based explanation suggests that knowledge-driven validation requires more processing resources than validating against prior text. In contrast, a task-based explanation suggests that the current findings were caused by differences in the strength of the inconsistencies in the current paradigm. Because the knowledge inaccuracies were outright errors whereas the text incongruencies were merely incongruent, the knowledge inaccuracies might have been stronger than the text incongruencies and therefore recruited a larger number of brain regions. Varying the strength of the inconsistencies can provide insight into which of these interpretations is accurate.

Another interesting difference between text-based and knowledge-based validation is that all regions that were involved in text-based monitoring were sensitive to *coherence* rather than *incoherence* of information with the text, whereas knowledge-based monitoring involved both regions that were more active for true information and regions that were more active for false information. A common assumption in coherence-monitoring research is that processing inconsistent information requires more effort and, thus, more resources than processing consistent information. Therefore, most neuroimaging studies focus on brain regions that are more active in response to incongruent than to congruent information. However, our results show that text-based monitoring only involves regions that are sensitive to *congruency* rather than incongruency of information. In itself, this congruency effect may not be a striking finding, as it is plausible that there are regions specifically involved in fitting congruent information into the existing mental representation (see also our discussion of the role of the right IFG) and some studies have also shown more activation for coherent than incoherent texts (Although these studies compared cohesion and coherence, which are both manipulations on a text-level, e.g., Ferstl & Von Cramon, 2001; Siebörger et al., 2007). However, what is surprising is that we did not find *any* brain regions that showed sensitivity to incongruency of information with the text, which suggests that processing an incongruent target does not recruit other brain regions than those that are already active in processing congruent information. Interestingly, the behavioral results did show incongruency effects in terms of reading times. Thus, it seems that processing incongruent information takes longer than congruent information, but it does not activate additional brain regions. We can speculate on possible explanations for this pattern of results. The extent to which incongruent texts are processed may depend on specific task demands. Prior work has shown that passive comprehension – compared to consistency judgment– prompts a processing approach that is focused primarily on establishing local rather than global coherence (e.g., Egidi & Caramazza, 2016). Hence, the passive reading task used in the present study may have decreased the effects of the contextual manipulation. However, the behavioral results show effects of both text and background knowledge, which suggests participants were still focused on establishing both local and global coherence. A second possibility is that the text incongruencies in the current study elicit a weaker neural response because they are not outright contradictions with the preceding text.

Taken together, our results provide evidence that there may be (partially) different cognitive mechanisms involved and that separate roles for contextual information and background knowledge should be assigned when describing the cognitive architecture of validation. However, there were also similarities in activation patterns between the two conditions, as well as interaction effects (e.g., in left IFG and precuneus), suggesting that readers integrate information from both sources to construct a coherent mental representation. The current study is an important first step in explicating the neurocognitive architecture of text-based and knowledge-based validation processes. Moreover, our results provide a fruitful base for constructing and testing more specific hypotheses about the interaction between the two systems.

Division of labor in the coherence monitoring network

Our second aim was to examine the neural correlates of text-based and knowledge-based monitoring by investigating the division of labor between specific regions that are known to play a key role in coherence monitoring. Furthermore, we aimed to identify additional shared and unique brain regions involved in either text-based or knowledge-based monitoring. Our ROI results showed different activation patterns in the dmPFC, precuneus and the left and right IFG, suggesting that these regions contribute differently to coherence monitoring processes. In addition to the pre-specified regions, the whole-brain results revealed activation in AF, MFG and HC during text-based validation and/or knowledge-based validation, indicating that these regions contribute to coherence monitoring as well. We will discuss the potential division of labor between these regions in more detail below.

Prior studies suggested a general role for the dmPFC in coherence-monitoring and integration processes (Ferstl et al., 2005; Hasson et al., 2007; Helder et al., 2017; Mason & Just, 2006). The findings in the current study allow for a more refined picture of the role of the dmPFC as the region appears to be particularly sensitive to false information (i.e., information that is inconsistent with readers' background knowledge). This finding complements research outside the reading domain showing that the dmPFC is involved in memory and decision making and is activated when an error is detected to signal to other brain regions that changes in cognitive control are needed to address the error (e.g., Ridderinkhof et al., 2004). Extrapolating to coherence monitoring during reading this would imply that the dmPFC may communicate to other brain regions that an inaccuracy is detected in the text and that additional processing is needed to resolve this inaccuracy (e.g., inhibition of inaccurate information or inference generation to assimilate mismatching information). The observation that the dmPFC is most sensitive to inconsistencies with background knowledge is compatible with this proposal, as such inconsistencies can be classified as outright 'errors' whereas inconsistencies with the preceding text are merely incongruent and thus may not evoke the same response. Note, however, that the thresholds for categorizing implausible input as actual errors could show variation across readers and situations as the judgments depend on a particular reader's background knowledge and self-calibrated evaluation criteria (Zacks & Ferstl, 2015).

For the precuneus we observed an interaction between text-based and knowledge-based monitoring, which is consistent with previous proposals that it is involved in coherence-building processes (Ferstl & Von Cramon, 2001, 2002; Kuperberg et al., 2006; Maguire et al., 1999; Mellet et al., 2002) and updating the mental representation of a text (Speer et al., 2009). However, whereas previous studies suggested that it is primarily sensitive to the incongruency of target information with prior text (e.g., Ferstl et al., 2008; Helder et al., 2017), we observed a more complex activation pattern. That is, the precuneus became particularly active when the target was either false, but part of a congruent text, or when the target was true, but incongruent with prior text. When the target matched or mismatched with both context *and* background knowledge, activation in the precuneus decreased significantly. This pattern is consistent with the idea that the precuneus becomes active when the situation model requires a major update (Speer et al., 2009). This is the case when the information provided by the context must be inhibited, adjusted, or overwritten to fit readers' knowledge of the world or when world-knowledge information becomes less reliable (or more ambiguous) due to prior text. Targets that matched with both context *and* background knowledge only required a minor update and therefore may not elicit activation in the precuneus. Furthermore, targets that mismatch with both context and background knowledge make little sense and may not elicit an update at all.

In addition, our study showed an interaction between text-based and knowledge-based monitoring in the left IFG. These results support the idea that the left IFG plays a general role in coherence building (Ferstl & Von Cramon, 2001, 2002; Kuperberg et al., 2006; Maguire et al., 1999; Mellet et al., 2002) and in processing world knowledge violations (Menenti et al., 2009). Our results, however, allow for a finer-grained account. Although the left IFG showed high activation for all target sentences, the activation was most prominent when readers encountered false targets in a congruent text. In other words, it is sensitive to world knowledge violations, but evaluates these violations in the context of the text and becomes particularly involved when the context presents a reason to assume that information that seems false relative to world knowledge should not be dismissed easily.

Although the dmPFC, precuneus and left IFG showed different activation patterns, a common denominator is that they all became more active in the case of a mismatch (i.e., the target was false and/or incongruent with text). Interestingly, the right IFG displayed a different activation pattern, with higher activation for targets that were *congruent* with prior text[9]. This finding suggests that the right IFG is involved in integrating sentences with the preceding discourse, but only when the target presents a plausible continuation of that discourse. Moreover, since no influence of background knowledge was observed, discourse integration in the right IFG seems to proceed without validating the current input against world-knowledge information, indicating that some integrative aspects of text-based monitoring occur independently of knowledge-based monitoring. This conclusion is similar, yet not identical, to a proposal put forward by Menenti et al. (2008). Similarly, they suggested that the right IFG is sensitive to the (in)congruency between a target and the discourse context. However, they assumed that long-term memory -including world-knowledge of the type being tested with the materials in the present study- affects how information of current linguistic input is integrated with information of prior text, while in the current study we did not observe any influence of background knowledge in the right IFG.

In addition to these ROI-based results, the whole-brain results revealed involvement of several additional regions, including the AG, MFG and HC. The right AG and the MFG both showed an interaction between text-based and knowledge-based monitoring. Previous work implicated the AG in inconsistency processing

---

[9] It may be relevant to note that the whole-brain results show a main effect of background knowledge in the right IFG (MNI coordinates 50, 22 12). However, this is a slightly different region than the ROI discussed here.

(e.g., Helder et al., 2017; Menenti et al., 2009; Moss & Schunn, 2015), semantic and episodic memory retrieval (Binder & Desai, 2011), and in sustaining activated representations to support cognition (Guerin & Miller, 2011). With respect to coherence monitoring, perhaps the AG is involved in situation-model updating (similar to the precuneus) or detecting or resolving contextual or semantic conflicts (e.g., Ye & Zhou, 2009). The MFG is part of the frontoparietal control network (Vincent et al., 2008) and has been implicated in monitoring and manipulating cognitive representations in general (e.g., Koechlin et al., 2003) and, more specifically, in coherence-break detection (Hasson et al., 2007; Mason & Just, 2006) and coherence-break resolution (Helder et al., 2017). Thus, the MFG may be involved in retaining and manipulating the mental representation.

The role of the HC was particularly sensitive to ‘correctness’ of information (i.e., increased activation for true versus false targets), whereas the dmPFC is sensitive to ‘incorrectness’ of information. To better understand this pattern of results, it may be relevant to look at neural models of schema formation and adjustment (van Kesteren et al., 2010, 2013; van Kesteren et al., 2012). Specifically, the Schema-Linked Interactions between Medial prefrontal and Medial temporal regions (SLIMM) model (van Kesteren et al., 2012) has identified HC-mPFC interactions as important for the influence of existing knowledge on encoding, consolidation, and retrieval processes. According to this model, the HC can be seen as a register that stores the links to several parts of a memory, which is particularly important for novel (rather than schematic) memories(van Kesteren et al., 2017). In contrast, the mPFC helps integrating new memories with the existing knowledgebase. Based on this framework one might predict that false information would elicit more HC activation because the existing knowledgebase would need to be adjusted, whereas true information would elicit more mPFC activation because the information could easily be integrated into the existing knowledge structures. Interestingly, the direction of the activation in the current study was the exact opposite to what the SLIMM model would predict: the HC was more active for true than for false targets, whereas the mPFC was more active for false than for true targets[10]. This pattern of activation may be explained by the task used in the present experiment, which differed crucially from the tasks often used in memory studies. In most studies investigating memory participants are instructed to remember the information, even if it is false according to their background knowledge. The HC would then play a role in adjusting existing knowledge structures or in building new ones altogether. In contrast, in the current study participants were instructed to read the texts attentively for comprehension, but not necessarily to encode them in memory. Therefore, if participants encountered false information, the mPFC may have

---

[10] Note that the current paradigm activated a more dorsal portion of the mPFC than the region identified by van Kesteren et al. (2010). In fact, a more ventral region in mPFC was deactivated across all four conditions.

signaled the HC that the information is false and knowledge structures should *not* be updated, resulting in a different activation pattern than during other memory studies. Thus, our findings suggest that the HC-mPFC interaction during memory encoding is strongly influenced by the goal of the task (e.g., the reading goal).

Together, our results indicate a division of labor for the regions of interest. The dmPFC is mostly oriented towards knowledge-based processing, whereas the right IFG is mostly involved in text-based processing. Furthermore, the two streams of information affect the precuneus and left IFG interactively. Based on the pattern of results it seems that the dmPFC and the left IFG are involved in different aspects of inconsistency detection, as the dmPFC seems to detect erroneous world-knowledge information and signal the HC that existing knowledge structures should not be updated, while the left IFG evaluates world knowledge violations in the context of the text. The precuneus may be involved in repair processes, as it becomes deactivated either when there is nothing to repair (entirely congruent) or when the target makes little sense and is perhaps impossible to repair (entirely incongruent). Whereas all aforementioned regions seem involved in validating or monitoring incoming information, both the HC and right IFG seem involved in ‘normal’ integration processes when there is no explicit task to store the information in long-term memory. Interpreted as such, our results not only indicate that text-based and knowledge-based processing mechanisms recruit both shared (precuneus, left IFG, AG) and unique (dmPFC, right IFG, HC) brain regions but also that integration (in the right IFG) and validation processes (in the dmPFC, precuneus, and left IFG) operate independently to some extent. This seems in line with the assumptions of recent theoretical frameworks on text comprehension where connections in a text are formed in an initial integration stage and then checked against long-term memory in a validation stage (e.g., O'Brien & Cook, 2016b). Note, however, that neither the fMRI nor the behavioral data can confirm (or disconfirm) this integration-precedes-validation order of processing because neither dependent variable has sufficient temporal resolution to disentangle the time courses of these mechanisms.

In addition, validation processes seem predominantly knowledge-driven, as we only found regions involved in the detection (dmPFC) and evaluation (left IFG) of knowledge-based inconsistencies, but no regions involved in text-based detection and/or repair processes. Only after a knowledge inaccuracy is detected this information seems to be evaluated in the context of the text (in the left IFG). This evaluation could be part of initial detection processes (as it could indicate that congruency with the context is evaluated) but it could also be part of knowledge-based repair processes (e.g., using contextual information to resolve the knowledge inaccuracy).

Effects of processing inconsistencies on the memory representation

Finally, our third aim was to investigate how processes during reading affect the memory representation of a text. Behavioral results show that mismatching information -either with the text or with background knowledge- is remembered less well than matching information. These findings support the account that information that mismatches pre-existing memory representations, i.e., the readers' background knowledge or the mental representation of the text, is remembered less well (e.g., Anderson, 1983; Johnson-Laird, 1983; Zwaan & Radvansky, 1998). From the current results it is not clear whether these findings solely reflect a retrieval problem. In fact, mismatching information may not have been encoded in memory in the first place. Because of potential differences during encoding, our neuroimaging analyses only included trials that were remembered correctly in the memory task. To explore encoding differences between different trial types, future research should compare brain activation between items that were remembered compared to forgotten items. Due to the small number of trials in some of the conditions, this was not possible in the current study.

To conclude, an important goal of reading is to learn from texts, hence it is crucial that we understand how processes during reading affect the (long-term) memory representation. Results suggest that existing models of memory encoding and consolidation (e.g., the SLIMM model by van Kesteren et al., 2013) can aid our understanding of the memory processes involved in reading and provide a fruitful base for developing neurocognitive models of learning from texts, as they seem to share underlying neural processes. In turn, studies of reading comprehension can inform memory research, as investigating memory for texts presents an ecologically valid situation for testing models of memory.

Conclusion

Constructing meaning from discourse is a fundamental human ability that is intertwined with virtually all cognitive processes, including learning, memory, perception, decision making and language processing. A core issue in psycho-linguistic research is when and how different sources of information (e.g., syntax, semantics, discourse, pragmatics) interact. The current study examined the complex interplay between recently acquired knowledge (from the text) and long-term knowledge (from memory) in constructing meaning from language. More specifically, we studied how these sources affect online processing and the (offline) memory representation. By combining theoretical models originating from the discourse comprehension literature with (more) specific predictions of neurocognitive models of memory, we aimed to further our understanding of the neural underpinnings of text comprehension. Results of the current study are relevant for models of sentence and discourse processing and, moreover, for our understanding of how we construct meaning in a broader context (for example from spoken or visual input).

SUPPLEMENTARY MATERIALS

Results of model without splitting correctly and incorrectly remembered items

We ran additional analyses where we also included the incorrectly remembered items. For the whole-brain analyses the overall pattern of results is very similar. The contrasts that showed significant clusters were very similar to our original analysis, only the clusters in the congruent > incongruent contrast did not reach significance. With respect to the ROI analyses the main differences in the activation patterns is in the incongruent false condition, as the activation in the other conditions does seem similar to our original analysis. It is difficult to ascertain what caused these differences, as there may be different underlying processes involved. However, this pattern of results does seem in line with the notion that misremembered items may be processed differently compared to the correctly remembered items, as the items that showed the largest discrepancy were also the ones that were remembered the least on the memory task.

Region-of-interest analyses

For the mean activation in the dmPFC we found a main effect of background knowledge ($F(1) = 10.50$, $p = 0.003$): False targets elicited more activation than true targets. We observed a trend in the effect of text ($F(1) = 3.71$, $p = 0.064$). We did not find a text x background knowledge interaction ($F(1) = 0.88$, $p = 0.357$) (Fig. 3.4). For the mean activation in the left IFG we found a main effect of background knowledge ($F(1) = 9.46$, $p = 0.004$): False targets elicited more activation than true targets. We did not find main effects of text ($F(1) = 1.60$, $p = 0.216$) or a text x background knowledge interaction ($F(1) = 0.45$, $p = 0.509$) (Fig. 3.4). For the mean activation in the right IFG we found a main effect of background knowledge ($F(1) = 11.22$, $p = 0.002$) and a text x background knowledge interaction ($F(1) = 5.54$, $p = 0.025$), but no significant effect of text ($F(1) = 0.09$, $p = 0.772$) (Fig. 3.4). The right IFG showed more activation for congruent than incongruent targets if targets reflected true world knowledge information, but more activation for incongruent than congruent targets if targets reflected false world knowledge information. For the mean activation in the precuneus we did not observe any significant effects (Fig. 3.4).

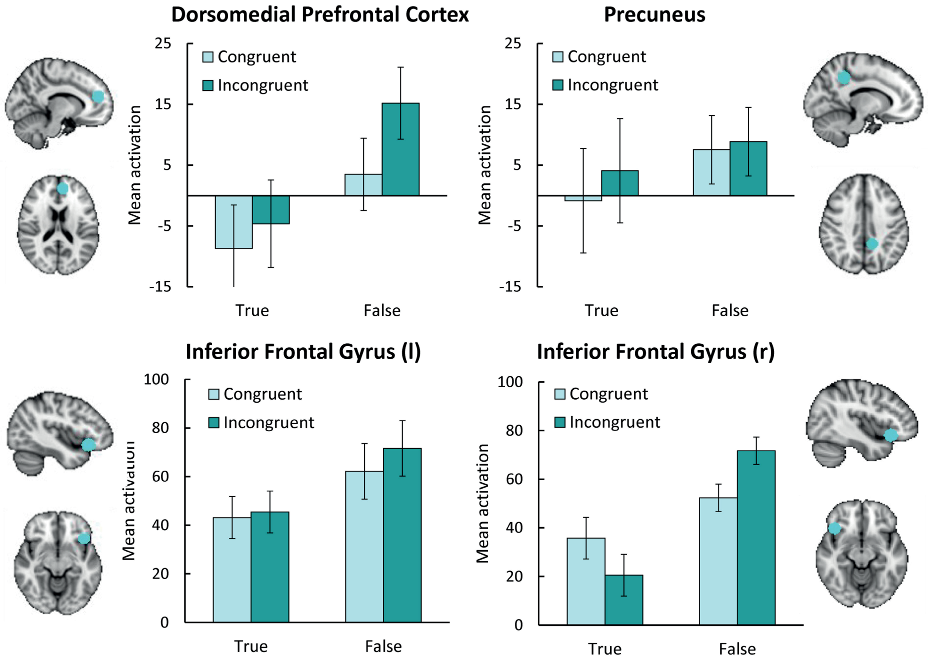

Figure 3.4. ROI mean activation in response to true/false and congruent/incongruent targets, for 8mm spherical ROIs centered at MNI coordinates [-9, 48, 21] (dmPFC), [-12, -45, 42] (precuneus), [-42, 24, -12] (left IFG), and [42, 24, -12] (right IFG). The left side of the brain is plotted on the right side of the image.

Whole-brain analyses of text-based and knowledge-based monitoring

We conducted a set of whole brain analyses to explore the involvement of other regions in text-based and knowledge-based monitoring. To examine the neural correlates of knowledge-based monitoring, we contrasted activation on trials in which participants read true targets with that on trials in which they read false targets and vice versa. The whole-brain contrast of false versus true targets resulted in clusters of activation in the right IFG, right superior frontal gyrus, right paracingulate gyrus, left posterior supramarginal gyrus, bilateral AG and bilateral caudate (Table 3.4; Fig. 3.5). The reverse contrast did not yield any significant clusters. To examine the neural correlates of text-based monitoring, we contrasted activation on trials in which participants read targets that were incongruent with the context to activation on trials in which they read congruent targets and vice versa. Both the whole-brain contrast of congruent versus incongruent targets and the opposite contrast did not yield any significant clusters of activation. The text by background knowledge interaction resulted in clusters of activation in the left IFG and bilateral middle frontal gyrus (MFG) (Table 3.4; Fig. 3.5). These regions showed more activation for incongruent than congruent targets if targets reflected true world knowledge information, but more activation for congruent than incongruent targets if targets reflected false world knowledge information. The reverse contrast did not yield any significant clusters.

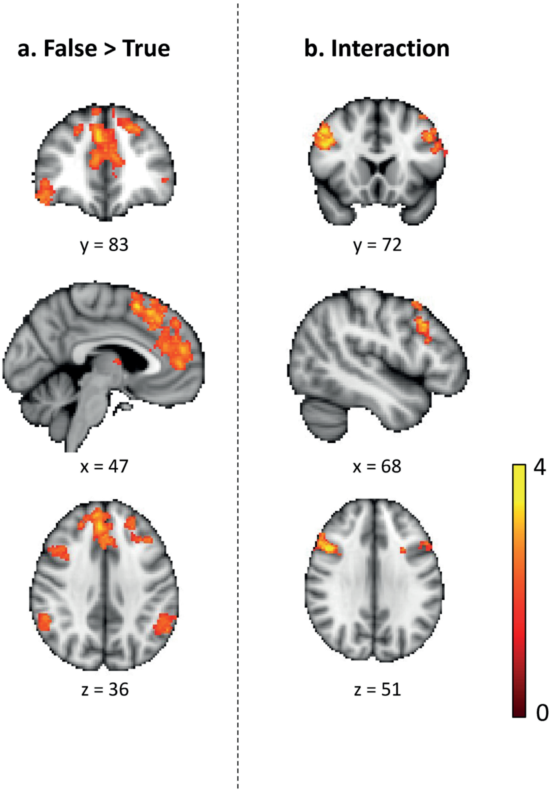

Figure 3.5. Whole-brain statistics maps for (a) the contrast false > true targets and (b) the interaction text x background knowledge across all participants (thresholded at $z = 2.3$ and $p < 0.05$). The left side of the brain is plotted on the right side of the image.

Table 3.4. Whole-brain group activations for targets in response to true/false and congruent/incongruent targets.

Anatomical region	L/R	Z max	MNI coordinates			voxels	*p*
			X	Y	Z		
a. Results for contrast false > true							
47% Superior Frontal Gyrus	R	4.23	14	14	60	4934	$p < 0.001$
77% Paracingulate Gyrus	R	4.18	8	48	10		
73% Paracingulate Gyrus	R	4.14	4	52	12		
41% Superior Frontal Gyrus	R	4.08	2	22	52		
55% Superior Frontal Gyrus	R	4.03	6	30	54		
47% Superior Frontal Gyrus	R	4.02	6	30	50		
32% Inferior Frontal Gyrus, pars triangularis	R	4.91	50	22	6	3099	$p < 0.001$
59% Inferior Frontal Gyrus, pars opercularis	R	4.5	56	18	8		
43% Frontal Orbital Cortex	R	4.21	50	26	-12		
45% Frontal Orbital Cortex	R	4.16	46	32	-12		
10% Inferior Frontal Gyrus, pars opercularis	R	4.07	58	20	-2		
71% Frontal Orbital Cortex	R	4.05	36	24	-20		
39% Insular Cortex	L	5.07	-32	22	-6	1301	$p < 0.001$
53% Frontal Operculum Cortex	L	4.03	-44	18	0		
37% Frontal Orbital Cortex	L	3.97	-28	12	-16		
52% Frontal Orbital Cortex	L	3.39	-38	22	-20		
12% Frontal Orbital Cortex	L	3.35	-52	24	-12		
21% Inferior Frontal Gyrus, pars triangularis	L	3.26	-54	22	-4		
41% Posterior Supramarginal Gyrus	L	4.45	-56	-50	26	569	0.011
33% Angular Gyrus	L	3.57	-52	-52	32		
40% Angular Gyrus	L	3.5	-54	-60	32		
48% Angular Gyrus	L	3.48	-50	-56	34		
54% Angular Gyrus	L	3.32	-56	-56	36		
41% Angular Gyrus	L	3.31	-48	-54	40		
98% Right Caudate	R	3.9	14	12	12	534	0.016
82% Right Caudate	R	3.52	14	2	16		
77% Left Cerebral White Matter	L	3.47	-10	6	0		
71% Right Pallidum	R	3.24	14	6	-2		
82% Left Caudate	L	3.12	-14	10	10		
89% Left Caudate	L	2.77	-12	0	16		

42% Angular Gyrus	R	3.58	58	-46	32	438	0.043
71% Angular Gyrus	R	3.57	60	-50	30		
72% Angular Gyrus	R	3.23	54	-52	40		
62% Angular Gyrus	R	3.11	52	-52	36		
48% Posterior Supramarginal Gyrus	R	3.04	54	-44	40		
26% Angular Gyrus	R	2.9	46	-52	16		
b. Results for contrast true > false							
No significant clusters							
c. Results for contrast incongruent > congruent							
No significant clusters							
d. Results for contrast congruent > incongruent							
No significant clusters							
e. Results for contrast interaction							
17% Inferior Frontal Gyrus, pars triangularis	L	3.51	-58	26	18	605	0.004
57% Middle Frontal Gyrus	L	3.19	-42	14	54		
36% Middle Frontal Gyrus	L	3.15	-36	8	44		
40% Inferior Frontal Gyrus, pars triangularis	L	3.14	-54	28	20		
55% Middle Frontal Gyrus	L	3.12	-42	12	48		
59% Middle Frontal Gyrus	L	3.08	-44	16	38		
41% Middle Frontal Gyrus	R	3.44	44	16	30	420	0.030
61% Middle Frontal Gyrus	R	3.34	50	20	40		
50% Middle Frontal Gyrus	R	3.26	52	22	34		
48% Middle Frontal Gyrus	R	3.13	50	18	34		
36% Middle Frontal Gyrus	R	3.01	54	22	30		
54% Middle Frontal Gyrus	R	2.76	52	30	30		
f. Results contrast negative interaction							
No significant clusters							