Adaptation to Semantic Violations of Varying Strengths Within and Across Texts

by

Michelle B. Colvin

B.S., Allegheny College, 2014

M.S., University of Pittsburgh, 2017

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Michelle B. Colvin

It was defended on

August 28, 2020

and approved by

Michael Walsh Dickey, Associate Professor, Department of Communication Science and Disorders

Julie Fiez, Professor, Department of Psychology

Charles Perfetti, Distinguished Professor, Department of Psychology

Dissertation Director: Tessa Warren Associate Professor, Department of Psychology

Adaptation to Semantic Violations of Varying Strengths Within and Across Texts

Michelle B. Colvin, PhD

University of Pittsburgh, 2020

Language comprehension is remarkable in that we adapt easily to different forms of language use, from adapting to speakers' dialects, meanings of new slang words, and fictional worlds described in novels. While there is growing evidence comprehenders adapt their expectations for text during reading, the nature of these adaptation mechanisms remains unclear. Some comprehension accounts maintain adaptation results from error-driven learning, such that larger errors lead to greater changes in expectations than smaller errors. Yet, the relationship between strength of error and the rate by which one adapts to semantic information during reading (i.e. whether larger errors lead to greater *and* faster adaptation) is an open question. The present dissertation tested this by investigating the time course by which younger and older adult readers adapt their expectations for fantasy-based texts. Participants read twenty short narratives, each containing five instances of semantic, or meaning, violations. Varying strengths of semantic violations–stronger (larger error) and weaker (smaller error) semantic violations–served as cues for readers to adopt a fantasy-world perspective on the text; adaptation was evident through decreased disruption to violations across instances within a single narrative and between narratives. The first experiment examined whether readers predict more fantasy-related content in subsequent parts of narratives with stronger than weaker violations during a cumulative cloze task. The second experiment used eye tracking to examine whether readers adapt faster to stronger than weaker violations within and across narratives. The third experiment took a broader cognitive approach to comprehension by investigating whether different aspects of readers' cognitive control ability associate with their degree of adaptation. Results indicate readers quickly adapt their expectations for a given fantasy text containing stronger violations. However, there was no evidence for adaptation to weak violations in fantasy narratives. There was considerably stronger evidence for adaptation to stronger violations within narratives than across narratives, suggesting there may be a limit to which comprehenders adjust their expectations during reading. Taken together, these findings are partially supportive of an error-based account of comprehension and leave open questions. This work highlights the importance of assessing comprehension from both language-specific and cognitive-general perspectives.

Table of Contents

List of Tables

List of Figures

Preface

Reading is the key that opens doors to many good things in life.

Reading shaped my dreams, and more reading helped make my dreams come true.

-RBG

1.0 Introduction

Psycholinguistic research these past two decades has made great strides in examining the concept of language as an adaptable versus static process. While once considered controversial, the idea that various sources of information (e.g. speaker dialect, amount of contextual support, etc.) present in language can each and combinatorially influence how we interpret and derive meaning has been important for informing theories of language comprehension. Furthermore, recent research has aimed at better understanding when and how these sources of information change in their utility for language processing across communication exchange. For example, comprehenders may first rely heavily on their general world knowledge to process language. Yet, they may adapt to rely on useful and elaborate contextual information later on during language processing once it is available. There has been substantial behavioral evidence that comprehenders adapt during both spoken and written language, and particularly growing evidence that readers adjust their expectations for text. However, it is less clear how the language system carries out such adaptation. The present dissertation addresses this question by investigating specific questions related to the nature of language mechanisms responsible for adaptation during reading comprehension.

The dissertation is organized as follows: I begin by providing a brief history of the ideas and frameworks that helped shape our current viewpoint of language comprehension as an adaptable process. Prominent early theories of language processing painted a static picture of comprehension, with little to no adaptation, or learning, occurring in such a system. Accruing evidence in favor of perspectives that allow learning and therefore change has led to multiple, current theories that incorporate change during comprehension. I describe two current comprehension frameworks and how each is able to potentially accomplish adaptation through its mechanism(s). An error-based account of comprehension (Kuperberg, 2016) is discussed in light of the broader error-driven learning literature that, together, provide specific predictions regarding the nature and time course of adaptation to different strengths of violating textual cues during reading. Three studies test these error-based predictions by investigating adaptation to semantic violations within fantasy-based text. Finally, the present work explores the additional question of whether adaptation mechanisms are guided by domain-general cognitive abilities of control.

1.1 Language Comprehension as an Adaptable Process

A long-standing theme of discussion in the history of language research has been the degree to which the language system is flexible. At the broadest level the flexibility, and thus adaptability, of language had been questioned through an old debate on whether language was a specialized system separate from the rest of cognition. Proponents on one side of the argument supported a view that language is special; a unique and innate language system separately stores information about sound and meaning. On the other side of the debate were researchers who adopted a more generalized cognitive view of the language system by assuming knowledge of various linguistic aspects of language was rooted in semantic memory, and this conceptual, semantic information was used similarly across cognitive systems.

This debate has played out in the field of language comprehension, specifically comprehension at the sentence level. Early linguists unequivocally argued that linguistic knowledge–broadly, specific knowledge about the constraints placed by the language–and world knowledge–more general knowledge regarding real-world probability–are represented separately

(Chomsky, 1965; Katz & Fodor, 1963). Processing theories supporting this specialized view of language propose that separability of these knowledge representations leads to timing differences in when each knowledge is available and used during comprehension. For example, modular theories of sentence processing (e.g. Frazier, 1987) purport that word-level representations of linguistic information are stored in specialized language modules (e.g. lexicon), that the comprehension system initially uses linguistic knowledge within these specialized modules first, and that only later does the system utilize separate general world knowledge. Hence, language comprehension at the sentence level was thought to be driven primarily by linguistic input and prioritized linguistic knowledge via a static, uniform bottom-up process. Crucial to this non-adaptive view of language is its implication that other sources of information present during language communication (e.g. contextual factors, characteristics of the language user) do not have much influence on the mechanisms of such a system.

Recent work has since challenged this non-adaptive framing of language and has pushed the field toward considering language as a more flexible process with respect to how information, including linguistic knowledge, general conceptual knowledge, and knowledge about the language environment, influences language processing. The first line of research has questioned the separability of semantic representations of linguistic and world knowledge. There is now experimental evidence for all sides of the debate: evidence in support of a strong separation between linguistic and world knowledge (e.g. Pylkkänen, Oliveri, & Smart, 2009), a weaker distinction between knowledge types (e.g. Willits, Amato, & MacDonald, 2015), and no distinction at all (Hagoort et al., 2004). While this literature contains flaws (Warren & Dickey, submitted), it still suggests that the roles of linguistic and world knowledge during processing are not as clear-cut as presumed in a modular account of comprehension. Another line of research has shown considerable evidence (Bicknell, Elman, Hare, McRae, & Kutas, 2010; McRae & Matsuki, 2009; Matsuki et al., 2011) for early effects of world knowledge on comprehension. From this body of work, it has been argued that both linguistic and world knowledge are available and can be used at any point in comprehension, including world knowledge driving early stages of processing (McRae & Matsuki, 2009). Yet another, and arguably the most compelling, line of research suggesting language as an adaptive process comes from evidence showing different sources of information are relied upon during processing across language tasks (e.g. Willits et al., 2015). This evidence implies the use of linguistic versus world knowledge during language processing is driven not by the availability of these kinds of knowledge but rather by the informativity or usefulness of the information given the task.

Effects observed in two studies from our lab (Milburn, Warren, & Dickey, 2015; Warren, Milburn, Patson, & Dickey, 2015) demonstrate differing informativity of, and thus reliance on, linguistic and world knowledge during comprehension. Warren et al. (2015) tracked eye movements as comprehenders read plausible sentences or sentences that were impossible due to semantic violations, or meaning violations. There were two types of semantic violations: violations of general world knowledge and selectional restriction violations (SRVs). Selectional restrictions, a form of linguistic knowledge (Chomsky, 1965; Katz & Fodor, 1963), were operationalized in the study as the basic set of semantic features a verb requires of its arguments. For example, the verb *jogged* has a selectional restriction for an animate object, and thus an inanimate agent jogging would be a SRV. Eye movements showed earlier and more disruption during processing of sentences with SRVs (e.g. *Corey's hamster entertained a nearby backpack and filled it with sawdust.*) than disruption during processing of sentences with world knowledge violations (e.g. *Corey's hamster lifted a nearby backpack and filled it with sawdust*.) and processing of plausible sentences (e.g. *Corey's hamster explored a nearby backpack and filled it with sawdust.).* In this violation study, linguistic knowledge regarding selectional restrictions is used earlier than world knowledge during processing.

Opposite results were found in a visual world eye-tracking study (Milburn, Warren, & Dickey; 2015) investigating linguistic and world knowledge's facilitative effects on language processing. Researchers directly compared facilitative processing effects associated with a combination of selectional restrictions and world knowledge of events to facilitative effects associated with only event knowledge. They did this by manipulating whether verbs in auditorily-presented sentences did (e.g. *Someone will pop the…*) or did not (e.g. *Someone will enjoy the…*) have selectional restrictions that combined with world knowledge activated by a visually-presented photo (e.g. photo of children with toys) to predict an upcoming object (e.g. balloon). Participants' eye movements to critical images in the photos measured anticipation for upcoming direct objects. Results revealed participants anticipated upcoming direct objects similarly for when facilitation was driven by event knowledge or a combination of selectional restrictions and event knowledge, suggesting that selectional restriction knowledge does not come into play earlier than world knowledge.

These seemingly contrastive results from Warren et al. (2015) and Milburn et al. (2015) reflect the usefulness of each type of knowledge for the two tasks (for discussion see Warren & Dickey, submitted). In Milburn et al. (2015), event knowledge was more useful because it facilitated a more precise prediction (e.g. balloon) than selectional restriction knowledge, which facilitated a category-level prediction (e.g. an object that is pop-able). In this sense, since selectional restriction knowledge did not contribute any unique information to the prediction (i.e. balloons are pop-able), event knowledge was the driver of facilitative processing. The utility of these knowledges is reversed for the processing of violations during reading. In Warren et al. (2015) selectional restriction violations were more informative that something had gone wrong during comprehension because they were more unexpected than violations of world knowledge. For example, it is more unexpected to encounter a *rock jogging* versus a *baby jogging*. Selectional restriction violations therefore resulted in a larger error signal during processing, which translated behaviorally into earlier and greater reading-time disruption.

It is now clear that comprehenders rely on sources of information that are most useful for a given task, and consequently at a particular time, language processing can be driven by general information about the world, language-specific information, context, or some combination of these. This raises the question of how an adaptive comprehension system is able to balance these sources of information. There are likely many types of knowledge present and this information can change in its informativity for a comprehender across a reading environment. Further, evidence from a noisy-channel account of language (e.g. Gibson, Bergen, & Piantadosi, 2013) suggests that a comprehender is able to switch and adapt their reliance from one source of knowledge to another when the original information becomes less informative for comprehension. What is the nature of the mechanism(s) that afford the comprehension system to be flexible and adapt (or not) to various influences of information? And what are the cues in a reading context that indicate one source of knowledge has become more useful for comprehension processes? One objective of this dissertation is to further investigate the patterns and extent to which semantic adaptation occurs during reading. Better characterizing cues of certain knowledge that are informative and serve as signals for a comprehender to update their beliefs/expectations (i.e. change reliance) can provide insight into the adaptation mechanism(s) that use such cues. One way to do this is by manipulating the strength of informative cues present in text in order to observe resulting adaptation behavior

(Studies 1 and 2). A second objective of this dissertation is to examine the cognitive underpinnings associated with such a language adaptation mechanism. Specifically, the present work takes an individual-differences approach to investigate whether an adaptation mechanism during language processing is associated with general abilities of cognition (i.e. our ability to control; Study 3).

1.2 Theoretical Accounts of Comprehension

Current theories of language comprehension (Cook & O'Brien, 2014; Kuperberg, 2016) assume comprehension is accomplished through a flexible, adaptive process. The RI-Val model of reading comprehension (Cook & O'Brien, 2014; O'Brien & Cook, 2016b) and the dynamic generative framework of event comprehension (Kuperberg, 2016) are two recent models of comprehension that come from different literatures and describe comprehension processes in very different ways. Yet, these models share the same underlying assumption that various types of information compete in the system to influence comprehension. Important to this dissertation, the models focus on different aspects relevant to the nature of adaptation mechanisms. The RI-Val model primarily provides a mechanistic account of the timing by which information becomes available and is utilized during processing, while the dynamic generative framework provides an account of how the informativity of information sources can evolve and change as new textual evidence is presented.

1.2.1 The RI-Val Model of Reading Comprehension

The RI-Val model takes a memory-based perspective on language comprehension (Gerrig & O'Brien, 2005; McKoon, Gerrig, & Greene, 1996; McKoon & Ratcliff, 1998) by positing that the processes primarily responsible for text-based discourse comprehension operate via passive activation. RI-Val joins other traditional activation plus integration theories of processing (e.g., Myers & O'Brien, 1998), as it can be considered an expanded model of the well-known Construction-Integration model (Kintsch, 1988). There are three stages of the RI-Val model. The first stage is the resonance (R) stage, in which a passive, dumb, and unrestricted activation process (Myers & O'Brien, 1998) automatically activates related information both from existing world knowledge in long-term memory and recently encoded contextual information in working memory. In this cascade of activation, reactivated memory representations serve as cues to activate other related representations in long-term memory. The result is a wide net of activated information, including information that is consistent and inconsistent with the text. Integration (I) is the next stage of RI-Val. During integration, activated concepts from the resonance stage are linked with textual information in working memory based on a general and superficial goodness-of-fit criterium. RI-Val is distinguished from previous activation plus integration models of text processing with its third separate stage of validation (Val). The validation stage is when initial linkages formed during integration are verified against the activated contents of long-term memory. A match occurs during this flexible, featural-level pattern-matching process if the linked concepts from the integration stage share many common features with information in long-term memory (for additional information, see Kamas & Reder, 1995).

The RI-Val model makes three critical assumptions regarding the way in which the three processing stages proceed during comprehension. The first assumption states that all stages of processing are passive. This is important because resonance, integration, and validation occur without and do not require strategic action from the reader. Once a process begins, it is said to run to completion, defined as a stable state (Kintsch, 1988). The second assumption states that the processes run in parallel but have asynchronous onsets, meaning that for integration to begin, at least two concepts from memory need to be activated via resonance. Similarly, initiation of validation only occurs once a minimum of one linkage is formed via integration. Crucial to this assumption, validation is believed to be dependent upon the output of the previous two stages: any factor that influences the timing of resonance or integration will also affect the timing of validation. This means that information more related to incoming content is more likely to be activated, integrated, and validated before information less related to incoming information. The third assumption is that once the reader reaches their coherence threshold (O'Brien & Cook, 2016a), they move on in the text. The coherence threshold is a subcomponent of standards of coherence (van den Broek, Risden, & Husbeye-Hartmann, 1995) and is considered a point in time at which validation results in a match. If a reader reaches their coherence threshold, they will shift their attention to subsequent information, regardless of whether the current processes have completed or not. This threshold for a criterion match can be set high or low by the reader during reading, depending on their goals for comprehension.

RI-Val describes a flexible and adaptive comprehension system through its assumption that contextual information and general world knowledge compete for influence over the course of the validation process. This assumption is particularly relevant to the current dissertation as it provides a distinct indicator by which adaptation can be observed. Competition in RI-Val can reveal when comprehenders change from previous reliance on one type of information to reliance on the other through the source of information dominating initial validation processes. At RI-Val's validation stage, initial linkages formed during integration are validated in a matching process against all contents in active memory, which includes both contextual information and general world knowledge. The ratio of contextual information to general world knowledge active in memory will determine what information is more available, and thus, heavily influences and dominates the matching process. For example, if a text contains weak contextual information, then resonance and integration will be driven by general world knowledge, and initial stages of validation will be dominated by this knowledge. If the text instead contains strong contextual information, then contextual information will be activated, integrated, and made available for initial validation more quickly than information from general world knowledge.

Recent studies from the Cook and O'Brien labs (Creer, Cook, & O'Brien, 2018; Walsh, Cook, & O'Brien, 2018; Williams, Cook, O'Brien, 2018) found experimental support for competition between context and general world knowledge within the RI-Val model. These studies utilized processing of semantic violations as a way to observe behavioral outcomes of this competition. This is the approach I take in the studies of this dissertation. As an example, Williams et al. (2018) manipulated text to include either weak or strong contextual support for semantic illusions (e.g. *Moses/Noah brought two animals of each kind on the ark*). They did this by altering the number of shared features (e.g. religious figure) mentioned in the text between the correct (e.g. *Noah*) and incorrect (e.g. *Moses*) terms. The logic is that strong contextual information for semantic illusions would lead early validation to detect a match between the illusion and shared features in the text, with a match indicating that processing proceeds smoothly. This is what researchers found. When there was a high amount of contextual support for shared features (6 features mentioned in the text), participants did not show processing difficulty until the spillover sentences, suggesting that strong contextual information dominated validation. Researchers attributed delayed comprehension difficulty in the spillover region due to RI-Val processes continuing and general world knowledge exerting influence on later validation processes. When contextual information was manipulated to be weaker, either by including weaker support for shared features (2 features mentioned in the text) in the match process of validation (Expt. 1) or by backgrounding contextual information so it was not readily available in active memory (Expt. 2), processing difficulty to semantic illusions in the text was observed early on in the target sentence containing the illusion. The researchers interpreted this result as initial validation being driven by general world knowledge.

The competition between general world knowledge and contextual information in RI-Val implies that the usefulness of these sources of information can change across a reading context. RI-Val processes run at each new textual input during reading. It is possible that general world knowledge could drive validation processes early on in a text, but as more information about the text is accumulated and becomes more useful for comprehension, validation processes later on during reading could then be primarily driven by contextual information. Results from Walsh et al. (2018, Expt. 2b) provide evidence for this pattern of adaptation within the context of RI-Val. The current dissertation follows this work by also examining the time point during processing when there is a change in reliance from world knowledge to contextual information. And it builds upon this work by further exploring the relationship between increased utility of information and reliance upon that information. It remains underspecified in the model how these sources of information change in their usefulness for incoming textual information across reading and drive adaptation. The resonance process of RI-Val is assumed to be a passive mechanism of memory retrieval. Any information that is relevant to the current textual input is activated, with only the most active concepts having an influence on comprehension. It is possible that increased usefulness of semantic information (i.e. stronger cue) in RI-Val reflects a change in the strength of the signal of that information during subsequent resonance processes. A greater signal during resonance would lead to an earlier and greater degree of activation in memory, which would then lead to quicker linkages during integration, and finally initial dominance during validation. Though it is difficult to imagine how a passive activation process becomes selectively more precise in its activation of only some information in short- and long-term memory.

1.2.2 The Dynamic Generative Framework of Event Comprehension

The RI-Val model is a theoretical framework describing how mechanisms rooted in memory accomplish comprehension via distinct stages. Information in long-term memory is activated, linked, and checked against textual information in working memory. The dynamic generative framework of event comprehension (Kuperberg, 2016) is another recent model of comprehension and explains comprehension in a much different way. Upon quick comparison, the dynamic generative framework describes competition between sources of information similar to that in RI-Val. Yet unlike the RI-Val model, this framework focuses on the way that expectations and beliefs about a text (e.g. reliances on information) are updated for subsequent processing based on current language processing. The framework's central focus on adaptation makes it an appropriate model to discuss regarding questions of this dissertation.

The dynamic generative framework of event comprehension (Kuperberg, 2016) builds upon constraint-based models from the sentence processing literature, models of the N400 from the ERP literature, and Bayesian models of language (Kuperberg & Jaeger, 2016). It posits that comprehension is driven by a reader's goal to infer the underlying cause of a given set of input. This underlying cause is the message-level interpretation, or event structure, that best explains the words encountered on the page. The dynamic generative framework describes a comprehension system by which multiple types of information–linguistic input, linguistic and semantic knowledge, and context–are weighted within a constraint satisfaction problem and, based on these weights, different information can more heavily influence comprehension (for review of work, Kuperberg & Jaeger, 2016).

The framework assumes a single mechanism processes all types of information. This mechanism includes a top–down activation component that is consistent with research suggesting that probabilistic preactivation plays an important role in language comprehension (Kuperberg & Jaeger, 2016). A reader builds a generative model that consists of a set of hierarchical representations from long-term memory. At the top of the hierarchical model are representations for the overall message of the text, or event structure representations, while at the bottom of the model are lexico-semantic representations for each word of the text. This generative model is the reader's best explanation for the available text at a given time. Crucial to the dynamic generative framework is the idea that this process is cyclical and constantly being updated. The reader holds a set of hypotheses for the event structure of a text at the top level of the generative model, and they actively generate probabilities of varying strength. A reader actively generates probabilistic predictions for semantic features of upcoming text according to these hypotheses. These predictions are then propagated from higher-level representations (e.g., of events, intentions, etc.) down to lower-level representations (e.g., of semantic features, sounds, and spellings). When new text becomes available, a reader learns whether their predictions are supported. What is not consistent with the predictions generates prediction error, which is defined as the difference between what was predicted and what was given in the text. Prediction error gets propagated through the generative model to update the reader's high- and low- level beliefs via Bayesian inference. This iterative cycle of belief updating and predictive pre-activation is what is believed to reduce prediction error across the model, and thus, allow the reader to converge on the most likely interpretation of the text.

There are two main assumptions to the dynamic generative framework. These assumptions are adopted from constraint-based approaches to sentence comprehension (e.g., Trueswell & Tanenhaus, 1994) and explain how the framework accounts for various sources of information. The first assumption is that initial language processing can be affected by multiple types of contextual information (e.g. semantic, syntactic). The strength of the evidence within the text determines to what degree each type of information influences processing. According to the model, competing context cues in the text are weighted for their reliability in predicting event structures. Only the most reliable cues or cue combinations influence what event is inferred. For example, strong and reliable contextual cues for a restaurant schema (e.g. *waitress* and *customer*) in the text would lead to high probability for a specific event (e.g. waitress served customer); the reader would activate semantic representations consistent with a word fulfilling that event (e.g. <serve>). The second assumption is that a reader is able to (though they may not always) use all sources of information available at a given time to build a higher-level representation of meaning. The framework assumes a reader completes the cycle of predictive preactivation and belief updating very quickly since natural reading is fast. Due to this time constraint on processing, a reader may infer an underlying event based on only some or all of the available textual evidence.

The dynamic generative framework clearly describes semantic adaptation during reading through its cyclical process of predictive pre-activation and belief updating. And it specifies the driver of adaptation as prediction error in the model. After a reader infers the mostly likely event structure of a text, they pre-activate semantic features for an upcoming word that align with event beliefs. Newly available text is compared with their predictions and any discrepancy leads to a certain amount of prediction error. If predicted semantic features partially share features with an encountered word, then this would result in smaller prediction error. Processing of the encountered word would be partially facilitated. If predicted features do not share many features with an encountered word, then this mismatch would result in larger prediction error, and disruption in processing to the word during reading. Prediction error drives adaptation by being the signal that gets sent back through the generative model via Bayesian inference to update the reader's beliefs. According to the model, the amount of prediction error directly correlates with the type of shift during belief updating. Small prediction error leads to a smaller shift, while larger prediction error leads to a larger shift from a previous high-certainty belief to a new high-certainty belief about another event. Put another way, large prediction error is more informative in the model for adaptation.

1.3 Error-driven Learning

The RI-Val model and the dynamic generative framework both describe comprehension as a flexible process: different sources of information can guide processing at various time points of a reading event. RI-Val specifies in detail the time course by which information compete for influence and inform initial processing. More useful and informative information wins out in the competition and influences early processing at each RI-Val stage. Yet the model only hints at how information might change in its informativity for comprehension processes. Such a mechanism would need to describe how textual cues become more useful over time and drive a shift from prior reliance on previously informative information to new reliance on these cues. Meanwhile, the dynamic generative framework proposes a specific adaptation mechanism through its belief updating/predictive pre-activation cycle. Prediction error is the driver of adaptation in the model. In addition, the framework assumes informativity of information operates at multiple levels of representation (Kuperberg & Jaeger, 2016). Prediction error gets sent back through the hierarchical model to update representations at the word level, structural level, event level, etc. RI-Val only considers the utility of information at word-level representations. Despite its encompassing explanation of how adaptation could occur, the dynamic generative framework is time agnostic with respect to its processes. For example, it is not as clear in the model how prediction error relates to the timing of adaptation (belief updating) in the cycle.

While the distinctions and shortcomings of the models differ, it has recently been argued that the theories can be seen as complementary (Colvin & Warren, 2020). It is possible that they focus on comprehension at different levels of explanation (Marr, 1982). The RI-Val model provides a description of comprehension processes at an algorithmic level of analysis. The model specifies memory representations and the processes for how these representations are used to carry out comprehension. Hence, RI-Val can be regarded as a mechanistic account. The dynamic generative framework is a Bayesian model of cognition, and many Bayesian models focus at Marr's computational level of analysis (Griffiths, Vul, & Sanborn, 2012). The framework describes comprehension as an automatic process during which readers continually generate and test hypotheses (Kuperberg, 2016). Through belief updating, the model explains a potential solution for comprehension. It is important to note that the dynamic generative framework does begin to push into the algorithmic level with its hypothesis for how a mismatch in semantic features generates prediction error.

The RI-Val model explains more of the *how* of comprehension, while the dynamic generative model focuses on the *why*. At the same time, RI-Val is underspecified as to how it accounts for adaptation during language comprehension. The dynamic generative framework provides a clearer framework for adaptation being accomplished by the comprehension system, but the specific details are lacking, particularly the nature of the relationship between prediction error, informativity of information, and timing of adaptation. So far adaptation during language comprehension has been discussed mainly in terms of semantic adaptation, or the process by which a reader changes their expectations for semantic information during comprehension. Broadening adaptation to include literature on error-based learning via syntactic adaptation as well as non-linguistic examples of error-based learning (i.e. adaptation to reward prediction error) may provide additional insight into the mechanism(s) that accomplish semantic adaptation during comprehension.

Many implicit learning and error-based models of language processing (Chang, Dell, & Bock, 2006; Chang, Janciauskas, & Fitz, 2012) have been informed by research on structural priming in production and comprehension. These computational models describe error-based learning in a simple recurrent network (Elman, 1990); error-based learning algorithms use the difference between a predicted output (expectation) and the correct output (actual) to adjust connection weights responsible for the prediction. The degree of learning that occurs is derived from the relative weights of connections. Such models posit that larger versus smaller prediction errors lead to more divergent weighting of preferences. In other words, a greater chance of an error leads to more learning. A surprisal term quantifies expectation probabilities in the models and is the main driver for changing these expectations during language processing.

Studies in the syntactic adaptation literature examining inverse frequency effects (Jaeger & Snider, 2008; 2013; Scheepers, 2003), effects in which a less expected prime structure is primed more strongly than a more expected prime structure, provide experimental evidence that larger prediction errors in comprehension lead to bigger changes in expectations than smaller prediction errors. In terms of surprisal, these results suggest larger surprisal for a less expected structure translates into a bigger increase in the probability of seeing that structure again (i.e. repetition). Hence, a less expected syntactic structure is a stronger–and more informative–cue for learning in this paradigm compared to a more expected structure. Further supporting an implicit learning account is the belief that inverse frequency effects are cumulative (see Jaeger & Snider, 2008 for discussion), meaning that estimating the probability of a structure involves tracking the distribution of many previous structures, not just the most preceding structure. The researchers argue that cumulativity allows for a better estimate of probability given the current context of language.

The dynamic generative framework is one form of an error-driven account of comprehension, and such, holds the similar assumption as other error-based models that larger prediction error leads to more learning/adaptation. Greater adaptation is described in the model as a greater shift during belief updating. In fact, prediction error is formalized as Bayesian surprise in the framework: the degree to which a reader shifts from their prior to their new probabilistic predictions following what is actually given in the text. If surprisal is the signal that carries informative information through the model to instill change, and there exists varying strength levels of surprisal (weak surprisal = smaller prediction error, strong surprisal = larger prediction error), then this implies 1) textual cues that generate prediction error (i.e. semantic violations) can be manipulated to be stronger versus weaker in a reading context, and from this manipulation, 2) the extent of and time course by which semantic adaptation occurs for the varying strengths of surprisal during comprehension can be observed. For example, manipulating cues to be stronger and more informative for a comprehender would result in larger prediction error/greater surprisal, which in turn, would lead to greater observed adaptation and more reliance on these cues for comprehension.

This still leaves the question of the timing by which adaptation might occur within the dynamic generative framework. Does larger prediction error lead to faster adaptation? Within the RI-Val model, more informative cues lead to quicker processing at each stage of comprehension. If RI-Val and the dynamic generative framework describe similar accounts of comprehension, and larger prediction error carries more informative information, then it is possible the answer is yes. Potential support for this hypothesis also comes from behavioral patterns of error-based learning in another cognitive system, specifically adaptation to rewards. Firing patterns of dopaminergic neurons to an unexpected reward mimic the teaching signal posited to underlie many models of reinforcement learning (e.g. Sutton & Barto, 1981), as these learning models state that learning about reward-predictive cues is regulated by prediction error. Reward prediction error signaling in the dopamine system suggests that larger signals of error produce more rapid updating of the cue value for rewards. For example, it has been shown that monkeys learn stimulus-reward associations faster in a high cost (stronger error) versus low cost (weaker error) learning environment (Tanaka, O'Doherty, & Sakagami, 2019). From these results, it is possible that within an error-driven account of language, larger prediction error during comprehension might lead to faster and greater semantic adaptation. Though not all cognitive, error-driven learning processes may work this way (e.g. motor learning, Takiyama, Hirashima, & Nozaki, 2015).

An important point to raise is that prediction within the dynamic generative framework (i.e. pre-activation of specific semantic features for an upcoming word in the text) is only one proposed computational explanation for how comprehenders anticipate upcoming content and adapt to new information. A different but related view is one that describes anticipation during reading as more a general state of preparedness for upcoming information rather than prediction of exact lexical upcoming input (for explanation, see Ferreira & Chantavarin, 2018). This view differs from prediction-based accounts (e.g. Kuperberg & Jaeger, 2016) that often pit integration and prediction against one another, and instead considers these processes as two sides of the same coin: integrative processes connect current input with previous context, while prediction (i.e. preparedness) aides processing of upcoming information by creating a state of greater receptivity to certain semantic features over others. According to this view, greater receptivity might occur only at the broad level of general ideas or events, at a semantic category level, or in some cases might be prediction through pre-activation of specific words. Because the concept of preparedness in this view is not entirely dependent on activation of lexical representations, it is less likely that adaptation processes would be sensitive to word-level probabilities. From this, there may be no expectation that adaptation patterns to stronger versus weaker signals would differ; a state of preparedness (versus lexical prediction) may allow a comprehender to be equally receptive to new information in both cases.

1.4 Adaptation to Semantic Violations

Thus far, I have considered theoretical explanations for how semantic adaptation might occur during reading. Error-driven learning accounts of comprehension such as the dynamic generative framework (Kuperberg, 2016) propose adaptation, or a change in a comprehender's belief state, is driven by surprisal/error in the model. Semantic violations present in fantasy text provide a good avenue by which to test the frameworks' predictions of surprisal and belief updating since these violations are unexpected in a reading context and can serve as informative cues to a comprehender to switch their perspective on the text in a specific way, from assuming the text occurs in the real-world to assuming it takes place in a fantasy world. Many of the existing fantasy studies in this literature have focused on the conditions under which general world knowledge and contextual information compete during processing (e.g. Creer, Cook, & O'Brien, 2018). The reasoning is that this type of perspective switch marks a shift from relying on general knowledge about the real world for language processing to relying more on contextual information about a fantasy world. One instance for which contextual information can drive a shift to taking a fantasy perspective is when textual information activates stored representations for a known fantasy world. For example, a person flying is a violation of our real-world knowledge, but Superman flying is a common event for those familiar with the comics. Experimental evidence has shown that comprehenders display no disruption when reading stories about known fantasy characters performing these types of violating events (Creer et al., 2018; Filik, 2008; Filik & Leuthold, 2008), suggesting that the name of a well-known fantasy character (e.g. Superman) is itself a strong enough cue to force comprehenders to quickly shift their perspective. Such cues are highly informative because they are tied to a wealth of semantic knowledge for the known fantasy world.

Belief updating, specifically switching to a fantasy perspective, has also been shown in studies that manipulate the amount of contextual information describing a novel fantasy world, one that does not have pre-existing, stored representations (Creer et al., 2018; Neiuwland & van Berkum, 2006; Walsh, Cook, & O'Brien, 2018). These studies are important because they suggest that there is a clear point during the competition of general world knowledge and fantasy contextual information at which textual cues become informative enough to force a perspective switch. Creer et al. (2018) showed that comprehenders relied on general world knowledge when reading stories about an unknown character's (e.g. Benjamin) fantastic events (e.g. bullets bouncing off Benjamin's chest) when no contextual information was provided prior to the actions. This was evidenced by strong disruption to the fantastic events. However, when the text included additional information at the beginning about Benjamin's powers, comprehenders relied more on fantasy contextual information during processing and no longer showed disruption to the events. Crucial to this dissertation, Nieuwland and van Berkum (2006) examined the time course of adaptation to specific violation cues within a single narrative. In this experiment, textual cues were strong violations of animacy, in which inanimate objects were main characters in the narrative who performed animate actions (e.g. a yacht talking at a therapy session). Results revealed a change in comprehenders' processing disruption to the animacy violations as the number of their encounters with the violations increased across a given narrative. The pattern was as followed: on the first few encounters with an animacy violation, readers exhibited processing disruption, but processing difficulty was decreased by the third encounter in the narrative and eliminated by the fifth encounter. Results from this study suggest that the process of perspective switching can be incremental. Multiple encounters with animacy violations became a strong signal for a comprehender to adopt, with high probability, a fantasy world belief.

Taken together, these studies in the fantasy text literature begin to address how the informativity of textual cues leads to perspective switching, particularly the strength of the cue necessary to trigger a switch. Yet, the mentioned studies only examined the time course of adaptation to cues that strongly violate general world knowledge, or cues that would result in large surprisal/prediction error. Creer et al. (2018) used well-known fantasy character names, while cues in Nieuwland and van Berkum (2006) were strong violations of animacy. Whether multiple encounters with weaker semantic violations during reading could also serve as informative cues for a perspective switch remains an open question. If so, this could provide further insight into the potential relationship between informativity, surprisal/prediction error, and belief updating (i.e. whether stronger violations lead to faster adaptation than weaker violations). The first aim of this dissertation addresses this question.

Another important issue regards the extent that comprehenders generalize their beliefs after a perspective switch. There is some evidence (Foy & Gerrig, 2014) suggesting that readers can extend prior knowledge associated with a known fantasy character (e.g. Shrek) to a novel character (e.g. Krum), particularly when the fantasy characters share a common link (e.g. Krum is Shrek's cousin). In this sense, the probability of cousins having the same characteristics is relatively high. Examining the extent of adaptation during reading has the potential to elucidate the types of probabilities generated and used by adaptation comprehension processes. Implicit learning accounts assume probabilities for upcoming text are based on cumulative information. Yet cumulativity has the potential to occur at multiple levels of comprehension. I have discussed how probabilities are generated from cumulative cues within a local reading context to force a perspective switch (i.e. animacy violations in a single narrative; Nieuwland & van Berkum, 2006). The second aim of this dissertation asks the question of whether probabilities are also generated based on the tracking of cues across a broader reading context (i.e. similar animacy violations across many narratives). If adaptation mechanisms interpret similar cues across narratives as cumulative, it is possible that readers might extend their fantasy world perspective they adopted for one narrative to their reading of all subsequent narratives.

1.5 Cognitive Control

The final aim of this dissertation is to investigate the cognitive underpinnings associated with semantic adaptation mechanisms. Language research presented throughout this introduction have provided insight into how comprehension processes are adaptable. Despite the flexible interplay between general world knowledge and contextual information (e.g. linguistic input) to influence comprehension, less is known about how we handle various informative and uninformative textual cues and disregard previous cues to rely on new informative ones as we update our beliefs. The fact that language communication is imperfect, yet readers are still able to effectively understand text (e.g. by ignoring parts of text, through reanalysis) indicates that the comprehension system involves some level of monitoring and control. In this sense, semantic adaptation can be considered one way (e.g. by updating beliefs) readers implement control during comprehension to overcome the imperfect of language, or when things go wrong while reading. Experimental tests of noisy channel accounts of language processing (e.g. Gibson, Bergen, & Piantadosi, 2013; Warren, Dickey, & Liburd, 2017) have begun to look at how comprehenders interpret semantic violations, by either considering them informative cues that the text is not set in the real world or simply considering them typos or errors present in the text. This is a good start as it tells us when cues are considered useful for adaptation. To add to this research, the current dissertation takes the approach of investigating the type of control necessary for comprehension mechanisms to carry out belief updating via perspective switching. I ask whether or not our general cognitive ability to control is related to our ability to adapt and switch perspectives.

There is considerable evidence that domain-general cognitive control processes are sometimes engaged during language comprehension (see Novick et al., 2010, for a recent review), and current research has focused on characterizing the role cognitive control plays in comprehension (e.g. Fedorenko, 2014). Cognitive control, also referred to as executive control, is broadly conceptualized as our ability to conduct higher-order processes in order to perform goal-oriented behaviors (Miyake, Friedman, Emerson, Witzki, Howertzer, & Wager, 2000). It is thought that cognitive control is comprised of two components, a conflict monitoring component and an intentional control component (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Conflict monitoring allows us to recognize when there are errors or conflicts that arise and to register the need for control to overcome these challenges, while intentional control is the actual act of changing behaviors in response for the need to control. There have been many distinguishable albeit related cognitive control processes identified in the literature. For example, a seminal study by Mikaye et al. (2000) found through factor analysis that three specific cognitive control processes were separable but shared a common underlying control factor. Mikaye and colleagues measured the following control processes: shifting, the ability to divide one's attention by shifting between tasks or mental sets; inhibition, the ability to control one's attention, thoughts, and behaviors in light of distractors; and working memory updating, the ability to mentally manipulate information such that new, relevant information is incorporated. As three core processes described in the cognitive control literature (e.g. Diamond, 2013; Mikaye et al., 2000), shifting, inhibition, and working memory updating are thought to differentially contribute to more complex cognitive control abilities, such as planning behaviors or problem-solving (e.g. conflict resolution).

Much of the investigation into the relationship between cognitive control and language comprehension has regarded conflict arising from ambiguity during processing (e.g. lexical ambiguity, Vuong & Martin, 2011; syntactic ambiguity, Hsu & Novick, 2016; Novick, Kan, Trueswell, & Thompson-Schill, 2009; Vuong & Martin, 2014). One such conflict is when a comprehender misinterprets an ambiguous sentence structure (e.g. garden path effect). Research suggests a causal relationship between general cognitive control ability and recovery from syntactic misinterpretation (Hsu & Novick, 2016; Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2014). For example, Novick et al. (2014) found that comprehenders who showed increased performance on a non-syntactic, cognitive-control training task had greater improvement in garden path recovery. It seems that language ambiguity resolution requires monitoring of ambiguous information and control over semantic/syntactic representations to override erroneous interpretations. Akin to this, adaptation to semantic violations via perspective switching might also be considered a form of conflict resolution. Semantic violations during reading are incongruent with a default, real-world perspective. Monitoring when this incongruence occurs may signal the need to control semantic representations in a particular way. One possibility could be that comprehenders shift between semantic representations for the real world and representations for fantasy worlds as they transition through a perspective switch. Another possibility could be that while semantic representations of a fantasy world dominate processing after a switch, representations of the real world must be inhibited to avoid interference. A third possibility is that working memory is updated such that fantasy world representations are incorporated in place of real-world representations, which are no longer considered relevant for comprehension. If perspective switching does require some level of general cognitive control ability, it seems more likely that some combination of shifting, inhibiting, and working memory updating processes may be at play.

1.6 Present Studies

The purpose of the present set of studies is to test specific theoretical questions regarding the nature of the mechanisms carrying out semantic adaptation during reading. The first question tests whether varying the severity of the error cue signaling adaptation leads to differences in the observed pattern of adaptation. Error-based learning accounts suggest that larger error leads to earlier, quicker adaptation than smaller error. The second question tests whether readers who switch perspectives in one text are more likely to make a similar switch in subsequent text. This will allow me to examine whether readers extend their updated beliefs to probabilities across multiple texts. The final question of this dissertation tests whether comprehension mechanisms responsible for semantic adaptation are correlated with the engagement of general cognitive control abilities. Cognitive control through inhibition, shifting, and working memory updating may have distinct or combinatorial influence(s) on a reader's ability to adapt to semantic violations.

Across all dissertation studies, I utilized processing of purposeful semantic violations present in fantasy-related text as a way to observe readers' semantic adaptation. Building from Nieuwland and van Berkum (2006), short fantasy narratives were created that described a main character performing human animate actions. Studies 1 and 2 address the first two questions of this dissertation. Semantic violations present in the narratives were manipulated to study whether the time course by which readers switch to a fantasy world perspective is moderated by the severity of the textual error cues. Stronger violations, denoting stronger error, consisted of inanimate objects performing human animate actions (e.g. hat owning a restaurant), following Nieuwland and van Berkum. Weaker violations, denoting weaker error, consisted of animals performing human animate actions (e.g. rabbit owning a restaurant). It was reasonable to consider an animal talking as a less severe violation than an object talking, as studies have suggested that the property of animacy is not binary but rather more graded in conceptual organization (e.g. in brain regions, Sha et al., 2015; in memory behavior, Loucks, Verrett, & Reise, 2020). Adaptation to both types of violations within a single narrative as well as adaptation across narratives was investigated. Study 3 addresses whether cognitive control plays an important role in perspective switching by taking an individual-differences approach. I examined whether there was an association between performance measures on cognitive control tasks of inhibition, shifting, and working memory updating and readers' pattern of adaptation to semantic violations in the narratives.

2.0 Experiment 1

Experiment 1 used a cumulative cloze task, in which participants read sentence fragments of short fantasy narratives containing semantic violations and were asked to produce what they felt was the most likely next action-word continuation of the stories. The current study was designed with a twofold purpose. The first was to examine how comprehenders adapt their expectations for text containing varying degrees of semantic violations. Participant responses in the cloze task functioned as indicators of readers' belief perspective of the text. The second purpose was for behavioral results in Experiment 1 to serve as norming data for Experiments 2 and 3. It was important to establish low cloze across all fantasy narratives so that any effects observed in Experiments 2 and 3 would not be confounded with predictability.

The current study tested whether processing of semantic violations of differing severity led to different patterns of semantic adaptation during reading. Semantic adaptation in the cloze task was operationalized as perspective switching–when a reader updated their expectations from a real-world perspective to a fantasy-world perspective. A reader's belief perspective of the text was observed via their produced words during the cloze task, meaning whether a reader produced action words aligning with a real-world perspective (e.g. a cow *mooing*) or a fantasy-world perspective (e.g. a cow *dancing*).

Error-based learning accounts of comprehension posit that the strength of error during processing determines the amount of signal to adapt. Within the dynamic generative framework, stronger prediction error leads to greater adaptation. To test this, short narratives were created that described a main character performing human animate actions. Semantic violations were designed such that they violated constraints placed by animacy similar to Nieuwland and van Berkum

(2006). To investigate whether the strength of semantic violations moderates the degree of adaptation during reading, I manipulated violations to be stronger versus weaker through the type of main character performing the human animate actions, with the logic being that stronger violations would result in larger error than that of weaker violations. There were three versions of each narrative: a stronger semantic violation condition in which an inanimate object agent performed human animate actions (e.g. a hat owning a restaurant; similar to Nieuwland and van Berkum), a weaker semantic violation condition in which an animal agent performed human animate actions (e.g. a rabbit owning a restaurant), and a control condition with no semantic violation in which a human agent performed human animate actions (e.g. a magician owning a restaurant). Each agent violation condition was compared separately to the control condition in order to measure adaptation.

The presence and general time course of adaptation to stronger and weaker violations was investigated within a given fantasy narrative and across narratives. It was predicted that participants would perspective switch during reading, such that they would produce animate action words following violating agents (inanimate objects, animals) as they moved through the cloze task. The RI-Val model of reading comprehension describes processing as being driven by the competition between contextual and world knowledge. If stronger violations serve as stronger contextual information in the cloze task, then participants may switch perspectives quicker in narratives with stronger violating agents than weaker violating agents. Finally, if readers extend their fantasy perspective belief such that they expect agents in all narratives to have human animate properties, then participants may produce animate-related actions words upon their first encounter with agents in subsequent narratives.

2.1 Methods

2.1.1 Participants

Forty undergraduates were recruited from the University of Pittsburgh psychology department participant pool. All participants were native English speakers and received course credit for participating.

2.1.2 Stimuli

Twenty short stories were constructed (see Table 1). Each story contained five sentences and described a main character carrying out human animate actions within a common schema. Main agents are in bold and critical action words are underlined in the example. These actions varied across stories whether they were performed individually (e.g. drinking coffee) or performed with other characters (e.g. having a conversation). There were five critical action words per story. The location of the critical words also varied per story; some stories had a critical word per sentence while other stories contained sentences with no critical word(s). There were three versions of each story for the three agent conditions: human (control condition) animal (weaker semantic violation), object (stronger semantic violation). The full set of stimuli can be found in Appendix A.

Table 1 Experiment 1 Example Narrative

There once was a **magician/rabbit/hat** that <u>owned</u> a new fancy restaurant downtown.
On opening night, the **magician/rabbit/hat** <u>invited</u> allhis friends, and he <u>advertised</u> his dishes to local food critics.
The **magician/rabbit/hat** was very <u>thrilled</u>, and the night went smoothly.
The **magician/rabbit/hat** <u>declared</u>, *Maybe this was a success*, as the last customer left the establishment.
This was the first time that the **magician/rabbit/hat** ventured in to the restaurant business.

Bold denotes agent type. Underlined denotes critical action word.

Agents across narrative conditions were chosen such that they related to one another. This was done by identifying 20 triplet sets of human-animal-inanimate objects (e.g. magician-rabbit-hat) that were semantically related via Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). The three possible pair-combinations (e.g. human-animal, human-inanimate object, animal-inanimate object) of the triplet sets of agents had on average an LSA value of 0.69. The triplet agent sets were then randomly paired with a constructed list of 20 unrelated narratives describing common schemas (e.g. owning a restaurant). This decision was made with the assumption that if it was unusual to find an agent in a particular schema, related agents would also be unusually found in the same schema.

2.1.3 Norming

Agent-verb combinations in the short stories were normed to verify they appropriately varied in their naturalness across the narrative conditions. Twenty undergraduates were recruited to rate the naturalness of the agent-verb combinations. These participants only took part in norming and did not complete Experiment 1. They received course credit for their participation. The agent-verb combinations described a triplet set–either a human, an animal, or an inanimate object–performing a human animate action (e.g. a magician/rabbit/hat *invited*). The agent-verb combinations were presented in a pseudorandomized order across three lists via Latin square so that participants only rated one version of an agent-verb combination with the same verb. Each participant rated 100 agent-verb combinations on a computer using a 7-point Likert scale (1= very unnatural, 4 = somewhat natural, 7= very natural) in Qualtrics. Participants were instructed to read the sentence fragment including the noun and verb and to rate how natural they thought it was for the noun to carry out the action of the verb. Participants saw each agent-verb combination one time and used the mouse to click on the number to make their response. After an initial round of norming, a small set of items did not show the intended pattern of naturalness, so additional combination sets were developed and normed with twelve additional participants to get a final set of items that included 300 unique agent-verb combinations (5 verbs to each triplet set, 20 stories total).

Naturalness ratings of the agent-verb combinations were compared to ensure that there were differing strengths of semantic violations across agent narrative conditions. It was expected that human agent-verb combinations would be rated as very natural, inanimate object agent-verb combinations would be rated as extremely unnatural, and animal agent-verb combinations would be rated somewhere in between. Norming supported this: combinations including a human were rated on average a 5.99 (SD = 0.72), combinations including an animal were rated a 2.93 (SD = 1.06), and combinations including an object were rated a 1.76 (SD = 0.67) on the 7-point scale. Separate t-tests comparing the different agent conditions confirmed that participants perceived semantic violations at different levels of severity: human-verb combinations (control) were rated as more natural than animal-verb combinations (weaker violation, $t = 27.17$, $p < 0.001$) and object-verb combinations (stronger violation, $t = 47.20$, $p < 0.001$), and, crucially, animal-verb combinations were rated as more natural than object-verb combinations ($t = 9.58$, $p < 0.001$).

2.1.4 Design

To examine adaptation to the different strengths of semantic violations, four counterbalanced lists were created via Latin square, such that all participants read narratives and completed the cloze task in the control condition and one semantic violation condition (either weaker violation condition or stronger violation condition). This between-subjects design ensured that participant responses for narratives making up each violation condition (weaker, stronger) were separately compared to responses for narratives in the baseline condition (no violation). This was important as it created clear comparisons to examine adaptation to stronger violations (stronger vs. no violation) and adaptation to weaker violations (weaker vs. no violation).

2.1.5 Procedure

Participants were instructed via an online survey through Qualtrics to read sentence fragments and, after each fragment, to provide what they considered to be the most likely next word. They read the fragments up to, but not including, the first critical verb or predicate adjective, and typed in a text box their response. Participants were told that after they responded, they would see how the author continued the paragraph. After clicking an arrow button, participants read the actual continuation of the story up to the next critical word and were again prompted to provide a typed response. Participants were instructed to take all the paragraph so far into account when guessing the next words in subsequent parts of the paragraph. This process continued until participants were able to read the whole narrative on the screen. Participants used the arrow button each time to advance through the stories and to advance to the next narrative. The experiment lasted between 15-20 minutes.

2.2 Results

2.2.1 Data Preparation

Participant responses for the cloze task were extracted from their narrative contexts. Responses ranged from one word to multi-word phrases. Only verbs and verb phrase responses were included in data analyses. About 6% of responses (e.g. adverb, noun, preposition, etc.) did not fit this criterion and were removed. The remaining verb responses were coded by two researchers (Tessa, Michelle) separately according to the following criteria: responses indicating plausible actions performed by only a human agent were coded a 1; responses indicating plausible actions performed by any animate agent (human or animal) were coded a 2; responses indicating plausible actions performed by any agent (animate, human and animal, or inanimate object) were coded a 3. Coded data by each researcher was compared and any discrepancies were discussed and resolved.

2.2.2 Statistical Analyses

Separate generalized linear mixed-effect (glmer) models were carried out using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R Studio (R Development Core Team, 2014; ver 3.1.1) to examine semantic adaptation to stronger violations and adaptation to weaker violations. Separate models were also run for adaptation within a given narrative and adaptation across narratives. The dependent variable for all models was participant responses. P values were generated using the lmertest package. Responses revealed that overall participants had low cloze (overall mean $\text{cloze}_{\text{average}} = 0.02$) across all completions in the short narratives. Crucially, low cloze was found across all agent narrative conditions: participants very rarely produced the actual next-word continuation for narratives involving a human agent (mean $cloze_{average}$ = 0.03), an animal agent (mean $cloze_{average}$ = 0.01), or an inanimate object agent (mean $cloze_{average}$ = 0.01). The full set of summary results for each model can be found in Appendix B. Here I only report any interaction effects as these results are pertinent to the present research question of adaptation.

2.2.2.1 Adaptation to Stronger Semantic Violations

Analyses examining adaptation to stronger violations present in the narratives compared cumulative cloze task responses for narratives containing human agents and responses for narratives containing inanimate object agents. It was predicted that processing of stronger violations would result in a large error signal if readers expected a continuation in the narrative consistent with an animate agent predicate, which would lead to greater adaptation, and that participants would adapt quickly if the size of the error signal drives the speed of adaptation.

To compare responses for human agent actions and object agent actions, coded participant responses were grouped as a binary outcome such that coded responses indicating actions performed by a human agent (code 1) and coded responses indicating actions performed by any animate agent (code 2) were combined into one group, and coded responses indicating actions performed by any animate or inanimate agent (code 3) constituted the other group. This decision was made with the effort to create the strongest baseline for comparison; in everyday language we commonly express humans performing both human-specific (e.g. boy talking) and animate-general (e.g. boy swimming) actions.

2.2.2.1.1 Within-narrative Adaptation

A model examining within-narrative adaptation included the fixed effects of agent violation condition (human, inanimate object), verb position (1-5), and the interaction of agent violation condition and verb position. Categorical fixed effect of agent violation condition was contrast-coded as the following: human = 0.5, inanimate object = -0.5, to reflect the prediction that participants would adapt to stronger semantic violations.

The maximally converging model included random slopes of agent violation condition and verb position by subject and agent violation condition by item. There was a significant interaction between agent violation condition (human vs. inanimate object) and verb position (1-5), beta estimate = 0.45, $p < 0.001$, indicating participants produced more animate-specific completions as they read narratives with inanimate object agents. Figure 1 displays the results. Upon visual inspection, adaptation within narratives seemed to occur very quickly, after the first exposure to a stronger semantic violation in the narrative.

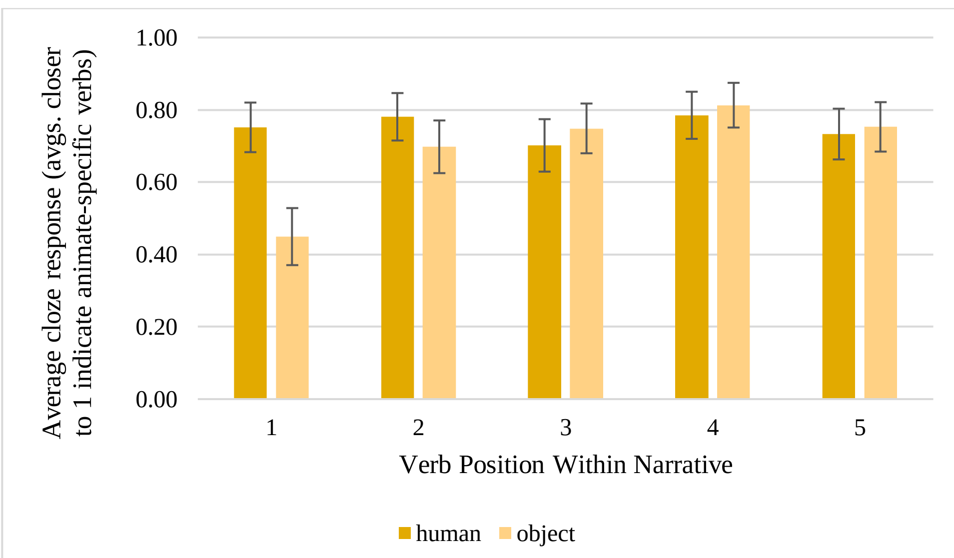


Figure 1 Adaptation to stronger violations within narratives. Comprehenders produced more animate-specific verbs upon their second encounter with an inanimate object agent in a narrative. Error bars represent standard error.

2.2.2.1.2 Across-narrative Adaptation

A model examining across-narrative adaptation included the fixed effects of agent violation condition and narrative position (1-20) and their interaction. Contrast-coding was identical to the within-narrative model, and the continuous measure of narrative position was scaled. Random intercepts of subject and item and random slopes of agent violation condition and narrative position were included to reach maximally converging models. Given results for within-narrative adaptation to stronger violations and to isolate across-narrative adaptation from any within-narrative adaptation effect, only the first completion for each narrative was included in analyses.

The maximally converging model included a random slope of agent violation condition by subject. There was a significant interaction between agent violation condition (human vs. inanimate object) and narrative position (1-20), beta estimate = 0.12, $p = 0.04$, suggesting participants adapted their beliefs for a fantasy-world perspective across narratives (see Figure 2). Upon visual inspection, the pattern of results suggest that participants had a strong shift across the first five narratives (1-5) and the next set of five narratives (6-10) to produce more action words attributing animate properties to inanimate objects. This suggests that adaptation may have occurred early on following the first few encounters with strong semantic violations.

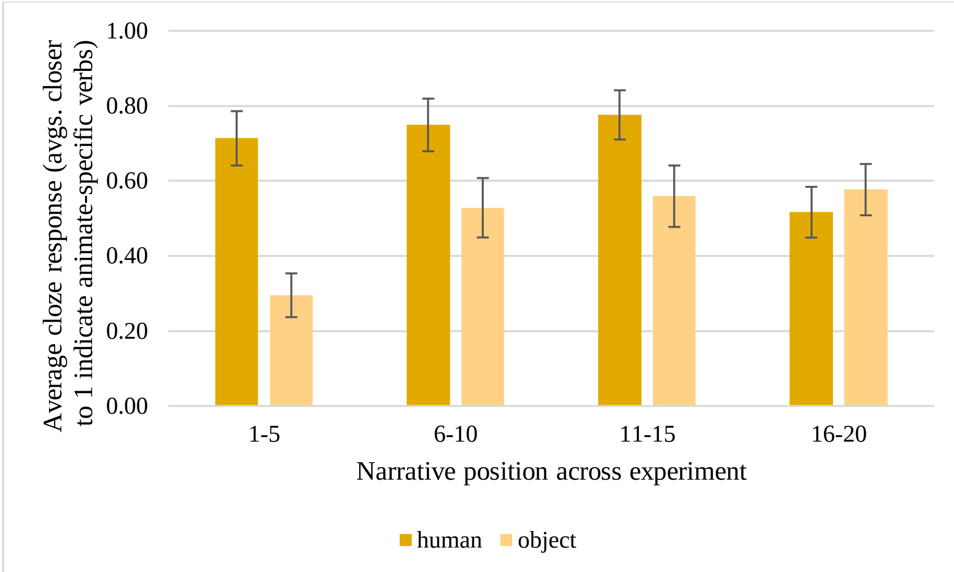


Figure 2 Adaptation to stronger violations across narratives. Comprehenders produced more animate-specific verbs upon their first encounter with an inanimate object agent in subsequent narratives across the experiment. For simplicity, narrative position (1-20) is displayed as four sections of the experiment. Error bars represent standard error.

2.2.2.2 Adaptation to Weaker Semantic Violations

Analyses examining adaptation to weaker violations present in the narratives compared cumulative cloze task responses for narratives containing human agents and responses for narratives containing animal agents. It was predicted that processing of weaker violations would result in a smaller error signal than that from processing stronger violations, and this smaller error would still lead to adaptation. However, comprehenders may take longer in the narrative to switch to a fantasy-world perspective following processing of weaker violations. To compare responses for human agent actions and animal agent actions, coded participant responses were grouped as a binary outcome such that coded responses indicating actions performed by a human agent (code 1) constituted one group, and coded responses indicating actions performed by any animate agent (code 2) or by any animate and inanimate agent (code 3) were combined into a second group. This grouping was performed so that a perspective switch could be clearly measured; comprehenders who switch to a fantasy-world perspective would produce human-specific actions for animal agents.

2.2.2.2.1 Within-narrative Adaptation

A model examining within-narrative adaptation included the fixed effects of agent violation condition (human, animal), verb position (1-5), and the interaction of agent violation condition and verb position. Categorical fixed effect of agent violation condition was contrast-coded as the following: human = 0.5, animal = -0.5, to reflect the prediction that participants would adapt to weaker semantic violations.

The maximally converging model included the random slopes of interaction of agent violation condition and verb position by subject and agent violation condition by item. There was a marginal interaction of agent violation condition (human vs. animal) and verb position (1-5), beta estimate = 0.16, $p = 0.08$. Visual inspection of the pattern of results suggests that the interaction is driven by the unusual result of more human-specific verb responses following animal agents than human agents in the first verb position of a given narrative (see Figure 3). Thus, the interaction does not suggest meaningful evidence of adaptation.

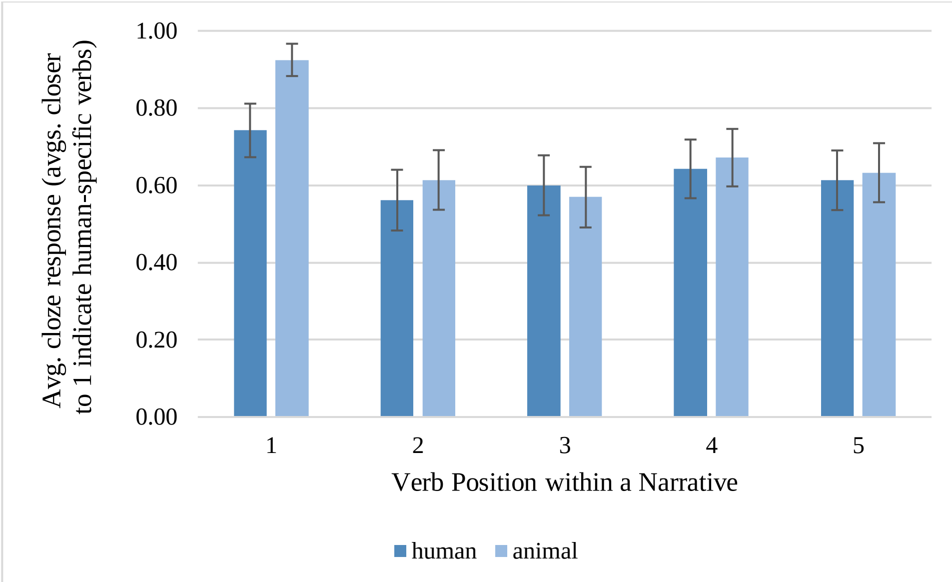


Figure 3 Adaptation trend to weaker violations within narratives. Comprehenders produced similar proportions of human-specific verbs following human and animal agents after their first encounter with a weaker, animal violation. Error bars represent standard error.

2.2.2.2.2 Across-narrative Adaptation

A model examining across-narrative adaptation included the fixed effects of agent violation condition and narrative position (1-20) and their interaction. Contrast-coding was identical to the within-narrative model, and the continuous measure of narrative position was scaled. Random intercepts of subject and item and random slopes of agent violation condition and narrative position were included to reach maximally converging models. Only the first verb completion for each narrative was included.

The converging model included only random intercepts of subject and item and no random slopes. There was no interaction between agent violation condition (human vs. animal) and narrative position (1-20). This lack of an interaction is attributed to an overall lower proportion of human-specific action responses for animal than human agents across the narratives.

2.3 Discussion

Experiment 1 used a cumulative cloze task to test whether the severity of semantic violations led to differences in the pattern of comprehenders' expectations for upcoming actions. Results suggest that behavioral patterns of adaption to stronger versus weaker violations did differ. There was clear evidence that participants switched to a fantasy world perspective upon processing stronger violations of animacy. This finding follows that found in Nieuwland and van Berkum (2006). Extending prior work, Experiment 1 also found evidence for adaptation to stronger violations across narratives, suggesting that participants adapted and expected all inanimate object agents in the narratives to have human capabilities. On the other side, there was no observed pattern of adaptation to weaker violations within and across narratives during reading. Cloze responses within narratives containing weaker violations were interesting in that they never showed initial disruption, as the proportion of responses indicating human-specific verbs following animal agents were actually higher than that following human agents upon the first encounter. Responses then remained similar across narrative conditions for subsequent verb positions.

Results from the strong violation narrative condition can be interpreted as potentially supporting an error-based learning account. Processing of stronger semantic violations produced large error, which led to a greater amount of adaptation. According to the dynamic generative framework, the amount of prediction error determines the strength of the signal in the model to update one's belief of the text. From the results, stronger violations of animacy served as informative cues for comprehenders to switch to a fantasy-world perspective, thus reducing prediction error across the model. Evidence for adaptation to stronger violations across narratives suggests that once comprehenders adopted a fantasy belief, probabilities aligning with this belief were generated for both individual inanimate object agents within a given narrative and the larger category of agent type across narratives. Furthermore, visual inspection of the time course by which comprehenders switched perspectives in the cloze task suggests that stronger violations may lead to quick perspective switching. This supports a key assumption of the RI-Val model that when stronger contextual information (i.e. stronger violation) is present in the text, it competes with and dominates real-world knowledge to influence processing. Because adaptation was not found in the weaker violation condition, the data cannot fully support an error-based account. Observed adaptation to both types of violations and a direct comparison of the violation strength types would be necessary to draw any conclusion whether perspective switching following stronger violations happens quicker than a switch following weaker violations, evidence in favor of error-based learning.

It is worth mentioning that the binary dependent variable of cloze responses had a different baseline group across models examining adaptation to stronger vs. weaker violations. Cloze responses for narratives containing object agents were grouped as a binary outcome according to whether they denoted actions plausibly completed by any animate agent (human and animal; baseline condition) or actions completed by any animate or inanimate agent (human, animal, and inanimate object). Cloze responses for narratives containing animal agents were grouped by whether they denoted actions completed by only a human agent (baseline condition) or actions completed by any animate or inanimate agent. These groupings of responses were appropriate to test our research questions because they allowed us to examine adaptation via measuring how often a participant produced a verb that did not typically fit with the given class category of agent (e.g. produced a verb requiring animacy for an inanimate agent). Despite this, it is possible that grouping decisions might have influenced the results found regarding adaptation to weaker violation within and across narratives. The grouping of cloze responses used to examine adaptation to weaker violations depended on the assumption that comprehenders provide mainly human-specific actions in narratives with human agents (the baseline condition). Yet, as mentioned this may not always be the case in everyday language. If comprehenders often chose general animate verbs in the human agent narratives, and this did not differ from the rate of general animate verbs produced for the animal agent narratives, then this would explain the lack of an interaction indicating adaptation. This seems to be the case for Study 1; the rate of overall responses produced indicating general animate actions (code 2) was almost equal at 46% of the total responses across human agent and animal agent narratives.

A separate objective of Experiment 1 was to establish that the sets of agent-verb combinations the narratives all had very low predictability. This was important as Experiments 2 and 3 used the same experimental items, and confirming low predictability in the cloze norming task eliminated any potential that adaptation effects observed during online reading in the following experiments could be confounded by high predictability of words or phrases.

In summary, findings from Experiment 1 showed that comprehenders adjusted their expectations following processing of semantic violations indicating inanimate objects or animals performing human actions, and that the pattern of adaptation differed for these semantic violations, with stronger violations of animacy leading to greater adaptation. The cloze task asked participants to overtly make predictions for upcoming text. Experiments 2 and 3 took the approach of measuring online reading behavior (eye movement measures, Exp. 2; reading time measures, Exp. 3) to test whether similar patterns of adaptation to stronger and weaker semantic violations would be found in more unconscious measures of language processing.

3.0 Experiment 2

Experiment 1 used a production task to test the nature of semantic adaptation to various violations during reading. Experiment 2 further investigated patterns of adaptation to varying degrees of semantic violations by having participants complete a comprehension task in which they read the same short narratives containing violations as in Experiment 1 while their eyes were tracked. The purpose of the current study was to examine the specific time course of processing by which comprehenders adapt to stronger versus weaker violations. Eye tracking allowed for this fine-grained investigation into the timing of adaptation by measuring comprehenders' disruption to semantic violations across reading. Results from previous reading studies looking at processing of violations (eye tracking, Filik, 2008; ERP, Nieuwland & van Berkum, 2006) support quick adaptation to strong violations. As previously mentioned in a study by Nieuwland and van Berkum, comprehenders showed initial disruption upon their first encounter with a strong semantic violation of animacy in a story but then attenuated disruption as they encountered additional violations, until disruption was eliminated by the end of the story. The current study tests whether the severity of semantic violations moderates the timing of adaptation by examining the patterns that comprehenders switch to a fantasy-world perspective following processing of stronger versus weaker violations. Furthermore, the study examined whether adaptation to stronger and weaker violations can be found not only within a given narrative, like Nieuwland and van Berkum, but also across narratives.

Adaptation via perspective switching in Experiment 2 was operationalized as a decrease in processing disruption to semantic violations across time. This can be explained as followed: once comprehenders switch to a fantasy-world perspective, processing of violations should become less difficult because beliefs now align with actions in the narrative. The dynamic generative framework posits that larger versus smaller prediction error signals greater updating of beliefs. If stronger violations produce more prediction error than weaker violations because they violate real-world beliefs to a greater degree (i.e. an inanimate vs. an animal performing a human animate action), then processing of stronger violations should serve as more informative cues to force comprehenders to update their beliefs about the text. Furthermore, if contextual information and real-world knowledge compete for influence during processing (RI-Val model), and stronger violations present in the text function as strong contextual information, then comprehenders should utilize this information during processing and adapt more quickly. Thus, it was predicted that comprehenders would show evidence of eye movement disruption upon their first encounter with both stronger and weaker violations, and that disruption to stronger violations would be attenuated quicker within a narrative and across narratives.

3.1 Methods

3.1.1 Participants

Participants were 76 undergraduate psychology students (47 female, mean age 18.87, range 18-24 years) who were recruited through the University of Pittsburgh subject pool (Table 2). Participants were native English speakers, at least 18 years or older, had normal or corrected normal vision, and were without any neurological impairment. Participants received one research credit for completion of the study.

Table 2 Participant Demographics for Experiment 2

		Total sample	Stronger Violations	Weaker Violations
Gender				
	n	76	38	38
	% Female	61.8	63.2	60.5
	% Male	38.2	36.8	39.5
Race				
	n	76	38	38
	% POC	22.4	18.4	26.3
	% White	77.6	81.6	73.7

3.1.2 Stimuli

The same 20 experimental narratives from Experiment 1 were used (Appendix A). The narratives were presented on the eye tracking computer monitor such that each of the five sentences of a narrative fit on a single presentation line with one blank line between each line of text. To accommodate this presentation, some narratives from Experiment 1 were shortened by removing one or two words (e.g. adjective) or by typing vs. writing out the full number (e.g. 100 pounds vs. hundred pounds). 10 additional filler stories and 2 practice stories of equal length were created (see Appendix C). The filler and practice stories included a human agent and were included so that participants were less likely to make strategic adjustments based on expectations of the content in the experiment.

Critical word regions of interest were verbs in the narratives and post-critical regions of interest included either a word or short word phrase following the critical verb, ranging from 8-16 character length. Critical regions are underlined and post-critical regions are in italics in Table 3. Yes/no comprehension questions followed half of the experimental and half of the filler stories. The answer to half of the comprehension questions was yes.

Table 3 Experiment 2 Example Narrative

Across the world, there was a diver/shark/cage that announced *all the baseball* games in his city.
At the start of a big match-up, the diver/shark/cage introduced *himself and* his partner to the crowd and then sung *their national* anthem.
After the anthem, the diver/shark/cage discussed *all the players* on each team and their statistics.
It was the fiftieth game of the season.
The weather became ugly, and the diver/shark/cage broadcasted *a delay of* game after the third inning.

Underlined denotes critical verb region. Italics denote post-critical region.

3.1.3 Design and Procedure

Participants' right eye movement and gaze location were recorded with an Eyelink 1000 tracker (SR Research Ltd., Toronto, Ontario, Canada) with a sampling rate of 1 ms. Participants viewed reading stimuli binocularly on a monitor approximately 63 cm from their eyes. Use of forehead and chin rests minimized head movements. The computer displayed all stories in a black Times New Roman font, size 20, against a white background. Before the experiment, the procedure was explained verbally, and participants were instructed to read normally and for comprehension. The experiment began with the instructions again in written form on the screen and a 13-point calibration. A single-point centrally located drift correction was performed after every trial. If the participants' apparent point of fixation was not within 2 degrees of the point, re-calibration was performed. A single fixation point in the upper left-hand corner of the screen ensured that participants were fixating just to the left of the first word of each narrative prior to stimulus display. Once participants completed reading each narrative, they fixated on a point to the bottom right of the narrative and pressed a key. Comprehension questions were displayed after 50% of trials. The experiment lasted between 20-30 minutes.

There were four constructed lists of stimuli for counterbalancing purposes. Participants read all experimental narratives but were presented only one of the two possible violating agent narrative conditions, such that participants only read narratives including human agents and animal agents or narratives including human agents and inanimate object agents. Stimuli were presented to participants in a pseudorandomized order so that no more than three of the same narrative condition were presented sequentially.

3.2 Results

3.2.1 Regions

Experimental narratives were divided into two regions of interest for analysis. The critical word region comprised the verb, and the post-critical region contained the following 1-2 words after the verb. There were five critical regions and five post-critical regions per narrative.

3.2.2 Preprocessing

Data from three participants were removed from analysis for low accuracy on the comprehension questions (more than two standard deviations from the mean average, < 70% accuracy). Three additional participants were removed from analysis due to poor tracking of the eyes throughout the experiment. Data from remaining participants (N=70) were used in the following analyses. Participants had on average 87% accuracy.

A procedure was performed on remaining data to pool short contiguous fixations. Fixations under 80 ms were combined with larger adjacent fixations on adjacent letters, whereas fixations under 80 ms and not within three letters of another fixation were eliminated. Prior to analyzing the data, full trials were removed if they had zero first-pass reading times on more than 50% of the trial (indicating track loss) and if they contained a greater number of blinks than number of fixations. Individual eye data points were removed within trials if during first pass there was a blink within a critical or post-critical region and the returning fixation did not occur within the same interest region or the next region beyond. Together this accounted for 4.9% of the data.

Three measures of reading behavior are reported. *Gaze duration* is the sum of all fixations in a region during first-pass reading until the eyes leave and move forward from the region. Gaze duration can include the initial fixation as well as any refixations in the region (prior to moving to the next region). *First-pass regressions out* is the percentage of trials that the reader has a regressive eye movement and leaves a region during first-pass reading. *Go-past* time is the duration of all fixations from first entering a region during first-pass reading until progression past that region. If there is no first-pass regression from a region, then go-past and gaze duration are equal. If there is a first-pass regression from a region, go-past includes all subsequent refixations on the region following the regressive movement until the reader moves on to the next region. All three measures are generally considered measures of early, first-pass processing (for descriptions, see Warren, 2011).

3.2.3 Statistical Analyses

Separate linear (lmer) and generalized (glmer) linear mixed-effect models were carried out using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R Studio (R Development Core Team, 2014; ver 3.6.2) for each of the three eye measures for critical and post-critical regions to examine adaptation to stronger violation and weaker violations within a narrative and across narratives. Here I report only a two-way interaction effect of experiment (stronger or weaker violation) and violation condition (violation or no violation) and three-way interaction effects of experiment, violation condition, and time (verb position for within-narrative adaptation; narrative position for across-narrative adaptation) of the models as these effects are indicative of disruption during processing and within- and across-narrative adaptation to the different violations during reading, respectively. P values were generated using the lmertest package.

3.2.3.1 Within-narrative Adaptation

Analyses examining within-narrative adaptation to stronger and weaker violations compared eye measures for critical and post-critical regions across a single narrative. It was predicted that comprehenders would adapt quicker to stronger violations within a given narrative. This would be shown through a quicker reduction in processing difficulty to stronger violations at either the critical or post-critical regions.

Planned models included the fixed effects of experiment (stronger, weaker), violation condition (violation, no violation), and verb position (1-5), and their two-way interactions and three-way interaction. In sum, there were no interaction effects found for within-narrative adaptation at the critical or post-critical regions. Full summary results from these initial models can be found in Appendix D, and select graphs from these analyses are included in Appendix E. The decision was made post-hoc to re-run models such that they investigated whether adaptation occurred between the first and second encounter with semantic violations in the narratives (versus previously investigating all encounters across the narrative, Neiuwland & van Berkum, 2006). These analyses were justified by observations in Experiment 1 suggesting that comprehenders switched to a fantasy-world perspective quickly in the narratives after their first encounter with a strong violation. These are the models reported below. They included fixed effects of experiment (stronger, weaker), violation condition (violation, no violation) and verb position (first, second), and their two-way interactions and three-way interaction. Categorical fixed effects of experiment were contrast-coded as stronger = -0.5, weaker = 0.5, violation condition as violation = 0.5, no violation = -0.5, and verb position as first = 0.5, second = -0.5, to reflect the prediction that stronger violations would yield greater and faster adaptation within a narrative. Dependent variables of the models were gaze duration (in milliseconds), binary first-pass regressions out (yes, no), and go-past (in milliseconds). The full set of summary results for the following models can be found in Appendix F.

3.2.3.1.1 Gaze Duration: Critical Region

The maximally converging lmer model included random slopes of verb position by subject and violation condition by item. There was no two-way interaction between experiment and violation condition. A three-way interaction between experiment, violation condition, and verb position, beta estimate = -68.175, $p = 0.031$, was found. Results revealed participants who read stronger violations in the narratives had greater reduction in their gaze duration upon their second encounter with the violations than participants who read weaker violations. See Figure 4 for results.

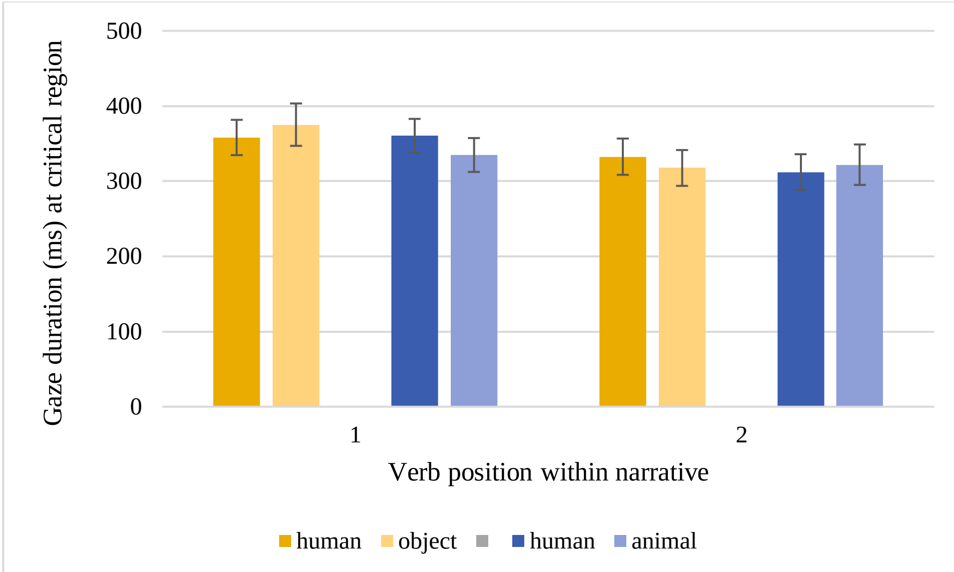


Figure 4 Fast adaptation to stronger (object) violations in gaze duration (ms) at the critical region within narratives. No evidence for adaptation to weaker (animal) violations. Error bars represent standard error.

3.2.3.1.2 Gaze Duration: Post-critical Region

There was no significant two-way interaction, nor a three-way interaction.

3.2.3.1.3 First-pass Regressions Out

There was no significant two-way interaction or three-way interaction at either the critical or post-critical region.

3.2.3.1.4 Go-past: Critical Region

There was no significant two-way interaction nor a three-way interaction.

3.2.3.1.5 Go-past: Post-critical Region

The maximally converging lmer model included random intercepts of subject and item and a random slope of verb position by subject. A significant two-way interaction between experiment and violation condition, beta estimate = -111.230, p = 0.014, revealed that participants spent longer reading (and re-reading) stronger semantic violations (*M* = 736.51 , *SD* = 794.36) and words following violations in the narratives than weaker semantic violations (*M* = 571.92 , *SD* = 447.87; see Figure 5). There was a significant three-way interaction: participants had a greater reduction in go past time at the post-critical region following their second encounter with stronger violations than weaker violations, beta estimate = -215.700, p = 0.018. See Figure 6 for the pattern of results.

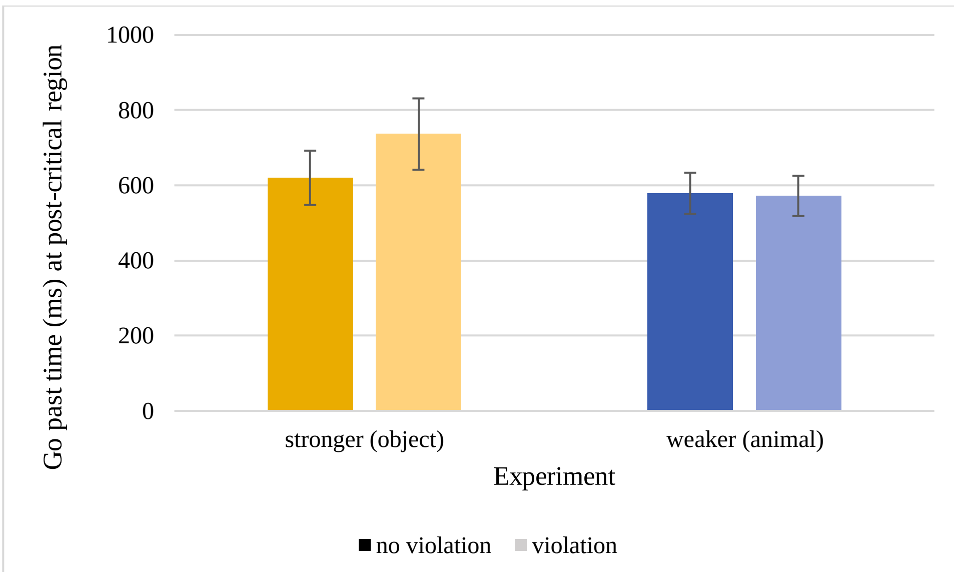


Figure 5 Longer go-past time (ms) for stronger (object) than weaker (animal) violations at the post-critical region. No disruption present for weaker violations. Error bars represent standard error.

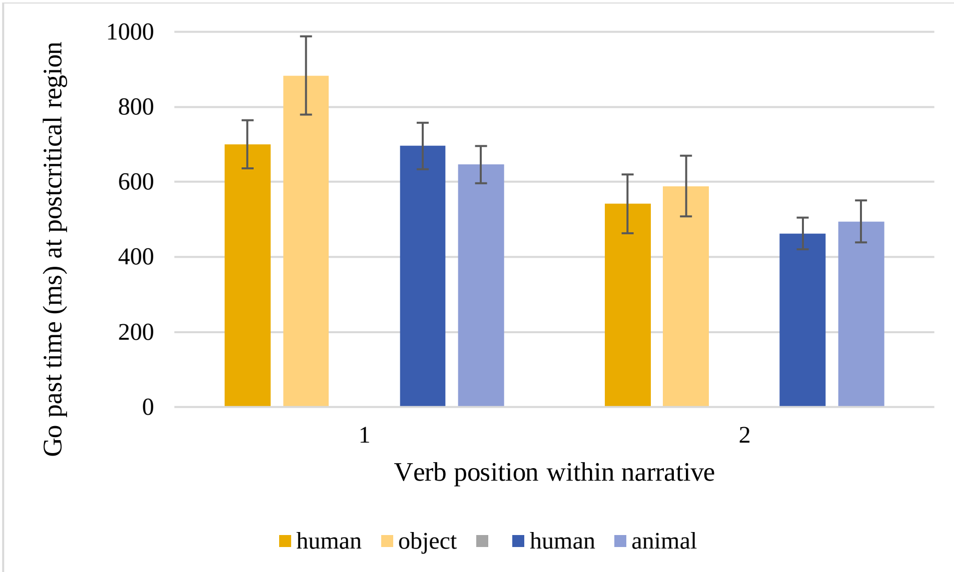


Figure 6 Fast adaptation to stronger (object) than weaker (animal) violations in go-past time at post-critical region No observed adaptation to weaker violation. Error bars represent standard error.

3.2.3.2 Across-narrative Adaptation

Analyses examining across-narrative adaptation to stronger and weaker violations compared eye measures for the first critical and post-critical regions across narratives. It was predicted that if comprehenders updated their belief for an agent in a narrative early on in the experiment, then they would also extend this belief to other similar agents across narratives. To examine this, participants' first encounters with violating agents in the first three presented narratives of the study were compared to the first encounters with violating agents in the last three presented narratives. Across-narrative adaptation was examined in this way so that any within-narrative adaptation effects present were isolated from across-narrative adaption effects observed.

Models included the fixed effects of experiment (stronger, weaker), violation condition (violation, no violation), and narrative position (first three, last three) and their two-way interactions and three-way interaction. Categorical fixed effects of experiment were contrast-coded as stronger = -0.5, weaker = 0.5, violation condition as violation = 0.5, no violation = -0.5, and narrative position as first three = 0.5, last three = -0.5. Dependent variables of the models were gaze duration (in milliseconds), binary first-pass regressions out (yes, no), and go-past (in milliseconds).

3.2.3.2.1 Gaze Duration: Critical Region

The maximally converging lmer model included random slopes of narrative position by subject and violation condition by item. There was a significant two-way interaction between experiment and violation condition, beta estimate = -62.789, p = 0.026. Participants had longer overall gaze duration on the critical violating word for narratives with inanimate object agents (*M* = 399.44 *SD* = 245.14) than narratives containing animal agents (*M* = 324.45, *SD* = 197.06). See Figure 7. There was no three-way interaction of experiment, violations condition, and narrative position.

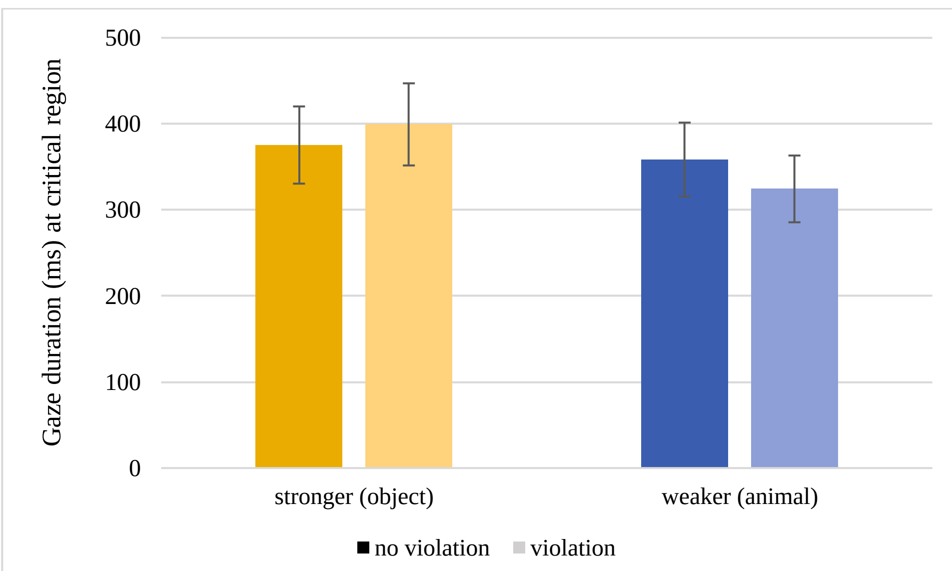


Figure 7 Longer gaze duration (ms) for stronger (object) than weaker (animal) violations at critical region. No disruption for weaker violations. Error bars represent standard error.

3.2.3.2.2 Gaze Duration: Post-critical Region

There was no two-way interaction or three-way interaction found for gaze duration in the post-critical region.

3.2.3.2.3 First-pass Regressions Out

A glmer model did not successfully converge for first-pass regressions out at the critical region. There was no two-way interaction or three-way interaction found for first-pass regressions out in the post-critical region.

3.2.3.2.4 Go-past: Critical Region

There was no two-way interaction or three-way interaction found for go past in the critical region.

3.2.3.2.5 Go-past: Post-critical Region

A two-way interaction of experiment and violation condition was found, beta estimate = -160.97, p = 0.046. Participants had longer go-past time on the post-critical region for narratives containing stronger violations (*M* = 897.71, *SD* = 871.51) than weaker violations (*M* = 647.55, *SD* = 413.92) See Figure 8. There was no three-way interaction.

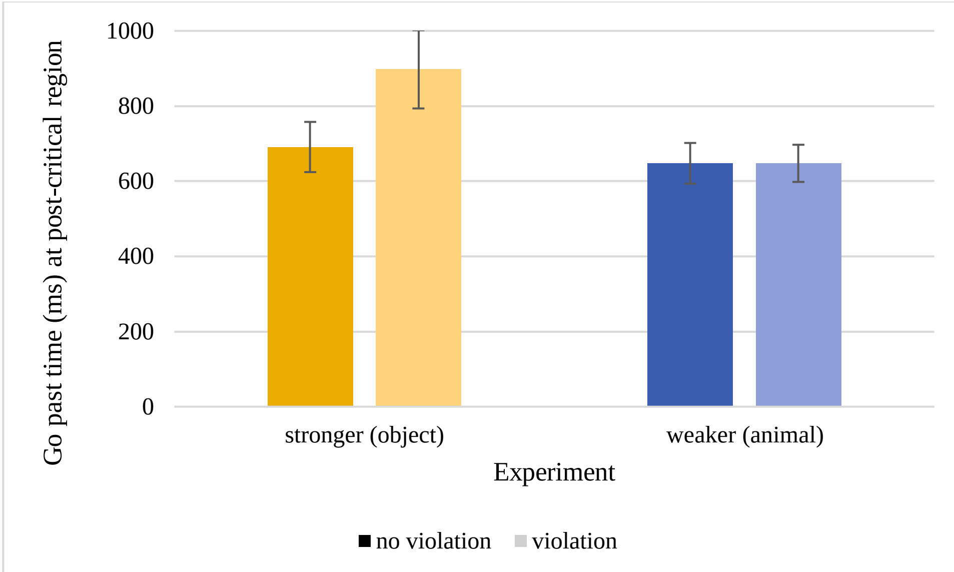


Figure 8 Longer go-past time (ms) for stronger (object) than weaker (animal) violations at post-critical region. No disruption to weaker violations. Error bars represent standard error.

3.3 Discussion

Experiment 2 tracked eye movements during a comprehension reading task to investigate the time course of adaptation to different strengths of semantic violations within a single narrative and across narratives. To see if participants adapted their beliefs in a single text, eye movement disruption to the first and second encounters with violations within the narratives were compared. Results supported the first hypothesis by showing participants adapted quickly to stronger violations after only a single encounter, and the adaptation pattern was significantly different than that of adaptation to weaker violations. Critically, evidence for fast adaptation to stronger violations was revealed in decreased gaze duration at the critical verb region and decreased go past time at the post-critical region. This suggests participants adapted their reading during first-pass processing, as gaze duration reflects readers' initial processing of a target word and go past likely reflects both lexical processing and integration processes (Rayner, Warren, Juhasz, & Liversedge, 2004; Warren, 2011). There was no evidence that comprehenders differed in their proportion of first-pass regressions out in the critical or post-critical region across narratives with stronger and weaker violations. Thus, the differentiated pattern of adaptation to strong violations was not so much driven by regressive eye movements, but rather the time spent processing the agent's action and previous text. Furthermore, the significant finding that comprehenders processed stronger and weaker violations differently across the narratives cannot be attributed to faster adaptation for stronger vs. weaker violations. Similar to Experiment 1, there was no evidence that comprehenders adapted to weaker violations in the narratives.

Evidence of fast adaptation to stronger violations within a narrative is congruent with error-based theories of comprehension that state larger error during current processing generates a greater and quicker change in expectations for future processing. The dynamic generative framework of event comprehension (Kuperberg, 2016) describes this as a cyclical process by which the difference between a predicted word and a given word, or the amount of prediction error in the model, drives the amount of updating of beliefs. Stronger violations present in narratives in the current study were created to be maximally conflicting with real-world expectations. Inanimate objects performing human animate actions violated expectations to a greater extent than animals performing those same actions. Thus, stronger violations served as stronger, more informative cues to comprehenders that their expectations for upcoming text were inaccurate, and this led to quick updating of their beliefs to a fantasy-world perspective, beliefs that accommodated actions such as a *hat or a rabbit owning a restaurant*. It is plausible that learning via belief updating could occur during early stages of processing, as semantic features for an upcoming word would reflect the new beliefs and would be pre-activated. Crucially, the study failed to find full support for an error-based account because there was a lack of evidence for adaptation to weaker violations.

There is a potential alternative explanation for the findings. The study was designed such that participants read either stories with only stronger violations or stories containing weaker violations. Because of this, baseline reading speed for the narratives was accounted for in the statistical models, but I was unable to account for participants' baseline disruption effects to stronger and weaker violations. Results showed that there was greater average disruption to participants' first encounter with stronger violations in each narrative than their first encounter with weaker violations. In fact, baseline effects to weaker violations were lower than reading effects observed for no violations, suggesting that comprehenders did not actually show any disruption evidence to animals performing human actions at any point in the narratives. It is possible that this difference might explain the three-way interaction effect: no disruption to weaker violations may mean that participants adapted so quickly to weaker violations on the first encounter itself in a narrative, and this very fast adaptation–single trial learning,–was not captured in analyses as an adaptation effect at the same magnitude as the pattern of adaptation to stronger violations. Put another way, if comprehenders more easily accepted an animal agent as a main character of a narrative based on their past experiences (e.g. animated cartoons like Tom and Jerry), then processing of weaker violations might have led to *faster* adaptation because there was a smaller difference in processing to overcome and beliefs could have been adjusted quickly and easily.

Discussing these findings in terms of the RI-Val model of reading comprehension (Cook & O'Brien, 2014) may provide insight into the cognitive mechanisms driving the timing of these effects. RI-Val holds the assumption that activation of contextual information and real-world knowledge compete in its earliest stage of resonance, and whichever source of information dominates activation initially influences its later stages of integration and validation. Thus, adaptation occurs in RI-Val when there is a shift in which source of information dominates initial processing. In the current study, comprehenders adapted their perspective by relying more on violating cues present in the narratives to guide processing, which would indicate a shift from real-world to fantasy-world beliefs. Results suggest participants who read narratives with stronger violations switched their perspectives quickly, as they showed decreased disruption to the violations upon their second encounter in their gaze duration and go past time. According to RI-Val, these findings can be interpreted as activated information in working memory now outweighing activated world knowledge in long-term memory to drive validation. Since the information in the narrative (i.e. semantic violation) *matched* activated contextual information (i.e. fantasy-related information), there was less disruption during processing, whereas, previous processing of text showed disruption because general world knowledge prevailed over working memory. Data from other studies (e.g. Filik, 2008; Warren, McConnell, & Rayner, 2008) have also shown similar patterns by which context can quickly mitigate disruption during and following the first encounter with a violation. Interestingly, participants who read narratives with weaker violations did not show the same decrease in disruption between their first and second encounters with violations in the critical and post-critical regions of a given narrative, and instead, showed an absence of any disruption to weaker violations at the beginning of the narratives. The RI-Val account would attribute a lack of disruption at these time points as instances when working memory was driving comprehension. Yet, this interpretation seems unlikely, especially at the first encounter with a weaker violation when contextual information was sparse. Importantly, another key assumption in RI-Val is that each stage overlaps during processing. In fact, many studies supporting the RI-Val model (for eye measures, see Cook, Walsh, Bills, Kircher, & O'Brien, 2018; for reading times, see Creer et al., 2018; Walsh et al., 2018; Williams et al., 2018) base their interpretation of results on timing differences of disruption during reading. The current study's absence of any disruption to weaker violations at the critical and post-critical regions does not necessarily mean that comprehenders did not eventually have processing difficulty associated with weaker violations; this disruption may have occurred further along in the sentence containing the violation.

The current study did not find evidence for adaptation across narratives. This was investigated by comparing the first encounters with violations in the first three narratives presented in the experiment with the first encounters with the last three presented narratives. It was believed that this comparison would provide the best possible chance of observing across-narrative adaptation because it would capture any early effects of belief updating in the beginning of the study. A lack of across-narrative adaptation did not support the hypothesis that participants would extend their updated beliefs for an agent in one narrative to similar agents in other narratives. This would suggest that participants did not complete the reading task with an expectation that all narratives would involve fantastical events, nor did they generate probabilities for agents in a broader semantic category, such as all inanimate objects. Further support for this interpretation comes from significant main effects of condition at the post-critical region in models examining first-pass regressions out and go past time, as well as a marginal main effect of condition at the critical region for go past time. This suggests that participants showed continued disruption upon their first encounters with stronger and weaker violations in each narrative across the experiment. There were clear patterns in the data that participants had overall longer gaze duration at the critical verb and longer go past time at the post-critical region for stronger violations than weaker violations. The disruption effect for gaze duration at the critical region is consistent with previous studies that found immediate disruption associated with semantic anomaly detection (e.g. Rayner et al., 2004), and corroborates the stronger violation manipulation in the study as being more severe.

Findings from Study 2 revealed that comprehenders adapted quickly to stronger violations during reading of short fantasy narratives. This pattern in the present eye tracking study resembles comprehenders' pattern of adaptation to stronger violations in the self-paced reading study (Experiment 1). Crucially, a lack of evidence for adaptation to weaker violations in Study 2 limits any conclusions that can be made regarding the dynamic generative framework and error-based learning accounts more generally. Further discussion of theories of comprehension can be found in the general discussion. The final experiment of this dissertation continues to investigate semantic adaptation by questioning whether the ability to perspective switch during reading is related to general cognitive control abilities.

4.0 Experiment 3

The first two studies tested specific questions related to the nature of comprehension mechanisms responsible for adaptation during reading. The final study of this dissertation takes a slightly different approach by examining the cognitive underpinnings of such mechanisms. In particular, the main aim of Study 3 is to test whether the mechanisms that carry out perspective switching during reading comprehension reflect the engagement of domain-general processes of control. To a degree, it may seem obvious that a certain level of control over cognition is required for comprehenders to be able to adapt their beliefs during reading. For example, comprehenders must be able to regulate their attention to informative cues in the reading environment which signal adaptation is necessary for successful comprehension. Otherwise, cues would be missed, and adaptation would not occur. It is of no debate that attentional processes are at play during language processing. Importantly, attention is thought to be a core underlying skill of cognitive control (see Chan, Shum, Toulopoulou, & Chen, 2008 for a review of attention in cognitive control models) and supports the ability to self-monitor behaviors (Stuss & Alexander, 2007). What remains unclear is how–once attended to–comprehenders implement control over textual cues to carry out adaptation. As mentioned, adaptation via perspective switching occurred in Study 1 and 2 when comprehenders switched from relying on real-world knowledge during processing to relying on contextual information containing semantic violations (i.e. cues for a fantasy-world perspective). Study 3 questions what kind of control might be required for this change in reliance to occur.

It is plausible, during a perspective switch, that comprehenders might implement a form of intentional control when transitioning from relying on semantic representations for a real-world perspective to relying on representations for a fantasy-world perspective. This transition may reflect an ability to *shift* from following one set of rules for certain information to following another set for other information, as a real-world perspective involves different probabilities than those of a fantasy-world perspective. During and following a switch, there is also the logical possibility that comprehenders implement control that affords their reliance on new textual cues and also allows them to suppress previous textual information that is no longer useful for their new beliefs. This form of control may involve the ability to *inhibit* previous information. And crucial to the assumption that comprehension relies upon both long-term memory representations and active information in working memory, it is possible that comprehenders implement control to incorporate new representations into their working memory at each iteration of language processing (e.g. each new textual input). Thus, this form of control may involve *working memory updating*. The three core subcomponents of cognitive control of *shifting*, *inhibiting*, and *working memory (WM) updating* have been studied extensively in the literature, and have been found to be separable albeit interrelated (Mikaye et al., 2000). In a seminal study by Mikaye and colleagues, participants completed a battery of cognitive control tasks, each specific to a particular core skill. Results from confirmatory factor analyses revealed that shifting, inhibiting, and WM updating were distinguishable yet also shared a common, unifying cognitive control factor. From this evidence-based dissociability, it is reasonable to think that one, some, or all core cognitive control abilities (i.e. common cognitive control factor) might be required for adaptation via perspective switching to occur.

To examine whether cognitive control abilities of shifting, inhibiting, and working memory updating are associated with the ability to perspective switch, it was important to find a population with a relatively wide range of performance on cognitive control tasks. Older adults with healthy aging comprise one such population with large variability in their cognitive control (see Braver &

Barch, 2002, for a theory of cognitive control on aging). The current study followed Experiment 2 by investigating perspective switching during reading of narratives that contained stronger and weaker violations. Older adults read the narratives in a self-paced reading task and completed a battery of cognitive control tasks indexing the abilities of shifting, inhibiting, and WM updating. It was predicted that there would be a relationship between cognitive control ability and perspective switching, such that individuals with higher performance across the core control abilities would show greater evidence of perspective switching during reading. It was also predicted that there would be an association between WM updating and perspective switching.

4.1 Methods

4.1.1 Participants

Participants were recruited through the Pitt+Me registry, a database of individuals within the Greater Pittsburgh Area. Participants were included in the study if they were a native English speaker, between the ages of 40-85, and had no history of a neurological, neuropsychological, or neuropsychiatric disorder. All participants were prescreened via email or phone prior to participation in the two sessions, and had normal or corrected normal vision, and passed a field of vision test (i.e. passable performance on freehand line bisection task) and hearing test (i.e. perceived 40dB at 0.5, 1, 2, and 4 kHz in at least one ear). Those who did not pass the screenings for either hearing and/or vision during the first session were excluded from the study and did not complete the second session. Due to the pandemic, data collection was cut short.

A total of twelve individuals participated in both sessions of the study. Participants (58.3% female) were between 43 and 79 years of age (*M* = 56.67; *SD* = 9.73). Individuals in the sample had between 12 and 20 years of education (*M* = 16.58; *SD* = 2.50). Time between the first and second sessions per participant ranged from two to three weeks. Table 4 provides individual participant characteristics.

Table 4 Participant Demographics for Experiment 3

	Total sample (%)	Average Education (years)
Gender		
n	12	16.6
Female	58.3	16.9
Male	41.7	16.2
Race		
n	12	16.6
POC	16.7	18.0
White	83.3	16.3

4.1.2 Stimuli and Design

4.1.2.1 Reading Task

The same 30 narratives (20 experiment, 10 filler) in Experiment 2 were used for the current study. Each of the five sentences of a narrative was presented on one presentation line with one blank line between each line of text. To accommodate this presentation, some narratives from Experiment 2 were shortened in their wording (for comparison, refer to Tables 2 and 3).

The narratives were presented in a self-paced reading paradigm, such that only one word or short phrase was visible on the computer screen at a given time and participants were tasked with reading at their own pace by pressing a button to advance through the narratives. Time between each button press indicated reading time on a given region. Critical word regions of interest were verbs in the narratives and post-critical regions of interest included a range of one to five words following the critical verb. The presentation segments for experimental narratives are marked with a pipe (|); the critical word region is underlined, and the post-critical region is in italics (Table 5). The full list of parsed stimuli can be found in Appendix G.

Table 5 Experiment 3 Example Narrative with Self-paced Reading Segments

Across the world, | there was a diver/shark/cage that | <u>announced</u> | *baseball games* | in his city.
Before a big match-up, | the diver/shark/cage | <u>introduced</u> | *himself* | *and then* | <u>sung</u> | *their anthem*.
After the anthem, | the diver/shark/cage | <u>discussed</u> | *each player* | and his statistics.
It was the fiftieth game | of the season.
The weather became ugly, | and the | diver/shark/cage | <u>broadcasted</u> |*a delay of game*.

Pipes (|) mark the beginning and end of each reading segment.

Experimental narratives were counterbalanced across four presentation lists using a Latin square design so that there was only one condition per item in each list, and each list included the control narrative condition (no violation) and one violating narrative condition (stronger, weaker violations). Narratives within a single list were pseudo-randomized so that participants did not read more than three of the same narrative condition in sequential order. Yes/no comprehension questions followed half of the short stories. Participants completed the reading task twice across two experimental sessions, such that they read narratives with stronger violations in one session and narratives with weaker violations in another session. The order in which participants were presented lists containing stronger or weaker violations across experimental sessions was counterbalanced.

4.1.2.2 EC Assessment

Experiment 3 took a theoretically-driven approach (based on Miyake et al., 2000) to cognitive control task selection and analysis. Simple tasks to measure the three core cognitive control processes of shifting, inhibiting, and WM updating were administered across the two experimental sessions. The tasks, which are believed to tap each cognitive control process separately, were selected and adapted from previous research with healthy younger adults (Miyake et al., 2000) and older adults with aphasia (Allen, Martin, & Martin, 2012; Simic, 2019). There were six cognitive control tasks in total: two measuring shifting (Trails A+B, Verbal 1-Back), two measuring inhibition (Spatial Stroop, Go No-Go), and two measuring WM updating (Cued Shifting, Keep-Track).

The dependent variables for each group of cognitive control tasks were similar in nature and held the same directionality. These measures are as follows: the dependent variable for shifting tasks was shift cost in seconds, with smaller shift cost (faster reaction time) indicating better performance, and the two dependent variables for inhibition tasks, accuracy and reaction time, was the effect of interference, with smaller interference effects–greater accuracy and faster reaction time –indicating better performance, and WM updating tasks was accuracy, with greater accuracy indicating better performance. Each cognitive control task and its dependent variable(s) is summarized in Table 6.

Each cognitive control task was administered in the same modality for each participant. The majority of tasks were administered via E-Prime v.2 software (Schneider, Eschman, & Zuccolotto, 2002) on a Dell desktop or laptop, except for the Trails A+B task, which was administered as a paper-and pencil task (Appendix H). The Keep Track task was administered via computer but included paper to assist with participant responses (Appendix I). Participants were seated approximately 40 cm from the computer screen and were instructed to use two fingers they are comfortable with to make button-press responses on the laptop keyboard. Cognitive control tasks were pseudo-randomized per experimental session, such that participants never completed two consecutive tasks measuring the same cognitive control process. For most participants, this was three cognitive control tasks, one per group, per session.

4.1.3 Procedure

The reading and cognitive control tasks were completed either at the University of Pittsburgh or in participants' homes in quiet, well-lit conditions. Each experimental session consisted of the reading task and at least three cognitive control tasks, lasting about an hour. Participants were given instructions and practice before starting each task. They were compensated $10 per session plus transportation costs for their participation. Participants were able and encouraged to take a break between tasks to avoid fatigue.

4.2 Results

Experimental narratives were divided into two regions of interest for analysis. The critical word region consisted of a verb, and the post-critical region contained 1-5 words following the verb. There were five critical regions and five post-critical regions per narrative (20 experimental narratives). Time spent reading each region indicated processing difficulty.

4.2.1 Preprocessing

Data from twelve participants were included in analyses. Participants had 86% average accuracy on comprehension questions during the reading task. Raw reaction time for inhibition tasks and speed data for shifting tasks were assessed for outliers. Only reaction times from correct response trials were used. Any trial responses beyond three standard deviations (SDs) from the mean for a given task were removed prior to analyses. Overall, only 1% of the total reaction time data (i.e. inhibition tasks) and about 1.6% of the total speed data (i.e. shifting tasks) were removed. There were no outliers in accuracy data for the Spatial Stroop inhibition task or in accuracy data for the WM updating tasks. To account for the unbalanced conditions in the study (i.e. 2x trials for the no violation condition vs. trials for the stronger and weaker violation conditions), half of the no violation condition trials from each participant were randomly removed prior to analysis, so that there as a total of 10 trials of each violation condition per participant.

Table 6 Cognitive Control Task Descriptions and Corresponding Dependent Variables, Grouped by EC Process

CC Process	Task	Modality	Description	Dependent Variable	# Trials
SHIFTING	Trails A+B	Paper and Pencil	Participants are timed using a phone app as they complete two tasks: Trails A entails connecting a series of numbered dots (1-15) in numerical order and Trails B, connecting an alternating series of dots in both numerical and alphabetical order simultaneously (e.g. 1-A, 2-B, 3-C). Errors were identified by the examiner and corrected by the participant immediately, adding to task completion time.	**Shift cost** *[Trails B - Trails A (seconds)]*	n/a
	Cued Shifting	E-Prime 2.0	Participants are cued with a written word (SHAPE or COLOR) presented on the screen, and must categorize the stimulus following the cue according to either its SHAPE (triangle or square) or COLOR (blue or yellow). Participants complete three blocks: two pure blocks (i.e. SHAPE only; COLOR only), and one mixed block (i.e. 4 SHAPE trials, 4 COLOR trials, 4 SHAPE trials, etc.). Participants are told in advance that the target cue will change every four trials in the mixed block.	**Global shift cost** *[Mixed block (mean RT) – Pure blocks (mean RT)]*	384 (128 per block)
INHIBITING	Go No-Go	E-Prime 2.0	Participants respond as quickly and accurately as possible to the presence of an X on the screen. In the “Go” condition, the arrow is flanked by asterisks (i.e. *X*). In the “No-Go” condition, the arrow is not flanked, and participants are instructed to withhold their responses.	**Commission errors** *[% responses on “No-Go” trials]*	100 (80 Go; 20 No-Go)
	Spatial Stroop	E-Prime 2.0	Participants view a single left- or right-pointing arrow appearing on the left, middle, or right sides of the screen and must respond to the direction of the arrow, while ignoring its location. The task has three condition: congruent (e.g. left-pointing arrow on the left side of screen), neutral (e.g. left-pointing arrow in middle of screen), or incongruent (e.g. left-pointing arrow on right side of screen).	**Interference Effect** *RT [Incongruent-Neutral Trials (msec)]* *Accuracy [Congruent-Incongruent Trials (%)]*	240 (80 per condition)
WM UPDATING	Keep Track	E-Prime 2.0 (presentation) Paper and Pencil (response)	Participants are prompted with the written words of two colors (e.g. GREEN, PURPLE) and then view a series of square color patches (red, blue, yellow, green, purple, grey) which appear consecutively in one of six locations on the screen. They must keep track of the last location of the two prompted colored squares and report the last location in which the target colors appeared.	**Accuracy** *[% correct responses]*	15 (10 stimuli views each trial)
	Verbal 1-Back	E-Prime 2.0	Participants view a continuous string of individually presented letters, and must indicate with a button press when the current letter is the same as the letter presented immediately (*n*-1) before it. No response is required on remaining trials.	**Accuracy** *[% hits + correct rejections]*	75

4.2.2 Composite Scores

Composite scores were developed for each cognitive control process of interest. Raw data from each task were first transformed into standardized z-scores, and then the two scores from tasks measuring the same cognitive control ability were averaged to create the composite score (for this approach in previous research, see Allen et al., 2012; Simic et al., 2019). Standardized scores for the *Trails A+B* and *Verbal 1-Back* tasks (reaction time data) were averaged to create a shifting composite for each participant. Standardized scores for the *Spatial Stroop* and *Go No-Go* tasks (accuracy data) were averaged to create the inhibition composite. And the WM updating composite was comprised of standardized scores from the *Cued Shifting* and *Keep-Track* tasks (accuracy data). Since reaction time data was only collected for one of the inhibition tasks (*Spatial Stroop*), a reaction time composite score was not created.

Smaller shifting and inhibition composite scores reflect better performance (i.e. smaller shift costs and interference effects). To ease interpretation, the sign of the composite score for WM updating was reversed to match the directionality of shifting and inhibition composite scores, such that smaller WM updating composite scores reflect better performance (i.e. higher accuracy). A single, combined composite score for overall cognitive control (CC) ability was created by averaging the standardized scores of all control tasks. Smaller common CC composite scores reflect overall better cognitive control processing. See Table 7 for individual participant cognitive control raw data and composite scores.

4.2.3 Correlations

Relationships among cognitive control task raw scores and composite scores were analyzed using Spearman's rho correlation coefficients. A summary of the correlations can be found in Table 8. Unexpectedly, there were no significant correlations among tasks measuring the same cognitive control ability. There were significant correlations between some individual inhibition tasks and WM updating tasks (i.e. Go No-Go and Keep Track, reaction time of Spatial Stroop and Keep Track; reaction time of Spatial Stroop and Verbal 1-Back). Not surprisingly, significant correlations emerged between shifting tasks and shifting composite scores, inhibiting tasks and inhibiting composite scores, and WM updating tasks and WM updating composite scores. Also, common CC was also significantly correlated with the shifting, inhibiting, and WM updating composite scores.

4.2.4 Statistical Analyses

Separate linear mixed-effect models were carried out in using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R Studio (R Development Core Team, 2014; ver. 3.6.2) for the critical and post-critical regions to examine adaptation to stronger and weaker violations within a narrative. Given the small sample of participants for individual difference analyses and in order to give each predictor the greatest chance of accounting for variance, separate models were run for each cognitive control ability to test for effects of and an interaction between adaptation and each composite score. I report here only two-way interaction effects of violation condition (stronger violation, weaker violation, no violation) and verb position (1-5), and three-way interaction effects of violation condition, verb position, and cognitive ability (composite score: shifting, inhibiting, WM updating, and common CC score) of each model as these effects indicate adaptation within a narrative and the relationship between adaptation and cognitive control ability, respectively. The full set of summary results for each model can be found in Appendix J. Violation condition was contrast-coded via two contrasts: a stronger violation contrast, stronger = 0.5 versus no violation condition = -0.5, and a weaker violation contrast, weaker = 0.5 versus no violation condition = -0.5. The dependent variable of reading time was log-transformed. P values were generated using the lmertest package.

Initial models without any individual-difference factors included were run at the critical and post-critical regions. The maximally converging models had an interaction of violation condition and verb position random structure by subject and violation condition random slope by item. At the critical region, there was a significant two-way interaction effect between weaker violations and verb position, estimated beta = 0.041, p = 0.036. However, this effect does not seem to be driven by adaptation to weaker violations. Instead, the inverse pattern occurred: average reading time to the first encounter with no violation (*M* = 944.42, *SD* = 391.71) was *greater* than that of reading time to a weaker violation (*M* = 812.33, *SD* = 300.74) Average reading times evened out by the second encounter in the narratives (see Figure 9). There was no significant interaction effect present for stronger violations and verb position. There were no interaction effects for either violation condition contrast at the post-critical region.

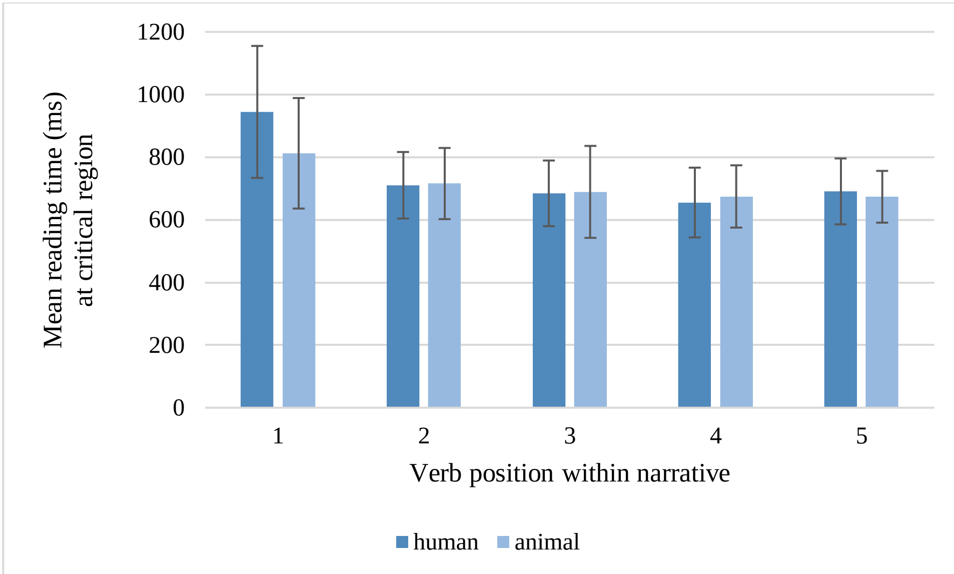


Figure 9 Greater mean reading times (ms) at the critical region for narratives with no violations (human) than narratives with weaker violations (animal). Error bars represent standard error.

Individual difference analyses were performed to examine whether variation in cognitive control ability predicted the magnitude of adaptation effects. It was predicted that comprehenders who had better overall cognitive control ability (common cognitive control, CC) would show greater adaptation to stronger and weaker violations during reading, and that WM updating ability would also predict the magnitude of adaptation, with greater WM updating scores predicting a quicker reduction in processing difficulty to stronger and weaker violations.

Table 7 Individual Participant CC Task Raw Scores and CC Composite Score

CC Process	Task	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	Mean	SD
SHIFING (sec)[2]	Trails A+B	14.55	59.50	22.87	47.77	26.09	20.00	6.74	12.16	11.31	24.92	11.41	62.58	26.66	19.30
	Cued Shifting	0.41	0.34	0.10	0.53	1.09	0.78	0.10	0.29	< 0.00	0.35	0.48	0.45	0.41	0.30
INHIBITING (Accuracy)[2]	Go No-Go	0.45	0.30	0.25	0.40	0.00	0.10	0.15	0.05	0.00	0.05	0.05	0.15	0.16	0.15
	Spatial Stroop	0.06	0.02	-0.01	0.00	0.01	0.04	0.01	0.04	0.00	-0.01	0.02	0.00	0.02	0.02
INHIBITING (RT, msec)[2]	Spatial Stroop	25.38	145.36	63.16	72.75	40.22	19.54	27.50	40.21	32.07	-58.41	28.30	43.52	39.97	46.12
WM UPDATING % Accuracy)[3]	Keep Track	0.93	0.70	0.50	0.67	0.97	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.89	0.17
	Verbal 1-Back	1.00	0.33	1.00	1.00	0.93	1.00	0.73	0.93	0.93	0.93	0.93	1.00	0.89	0.18

Composite Scores[1]

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Shifting[2]	-0.30	0.73	-0.61	0.75	1.11	0.44	-1.02	-0.58	-1.08	-0.14	-0.29	0.99
Inhibiting (Accuracy)[2]	1.98	0.65	-0.34	0.45	-0.61	0.27	-0.11	0.11	-0.88	-0.99	-0.17	-0.37
WM Updating[3]	-0.42	2.08	0.86	0.37	-0.34	-0.61	0.12	-0.43	-0.43	-0.43	-0.43	-0.32
Common CC[2,4]	0.42	1.15	-0.03	0.52	0.06	0.03	-0.34	-0.30	-0.80	-0.52	-0.30	0.10

[1]Composite scores were calculated by averaging standardized (z) scores of the raw data for each cognitive control process.
[2]Higher scores indicate worse performance.
[3]Signs of the WM Updating composite were reversed to match the directionality of the shifting and inhibition composites; higher scores indicate worse performance.
[4]Shifting scores, Inhibition accuracy, and WM updating scores were used to calculate the common CC composite.

Table 8 Spearman's Rho Correlation Coefficients for Individual CC Task Raw Scores and CC Composite Scores

		SHIFTING			INHIBITING (Accuracy)				WM UPDATING		
		1	**2**	**Composite**	**3**	**4**	**Composite**	**5**	**6**	**7**	**Composite**
1	Trails A+B	-									
2	Cued Shifting	0.20	-								
	Shifting Composite	**0.77****	**0.77****	-							
3	Go No-Go	0.38	-0.12	0.17		-					
4	Spatial Stroop (Acc)	-0.26	0.25	-0.01	0.28	-					
	Inhibition (Acc) Composite	0.08	0.08	0.10	**0.80****	**0.80****	-				
5	Spatial Stroop (RT)	0.53	-0.04	0.32	0.45	0.13	0.36	-			
6	Keep Track	-0.49	0.19	-0.19	-0.61*	0.35	-0.17	-0.62*	-		
7	Verbal 1-Back	-0.33	0.19	-0.09	-0.14	-0.08	-0.14	-0.61*	0.17	-	
	WM Updating Composite	0.54	-0.25	0.19	0.48	-0.17	0.20	0.80**	**-0.75****	**-0.78****	-
	Common CC Composite	0.69*	0.30	**0.64***	0.73**	0.32	**0.66***	0.74**	-0.56	-0.50	**0.69***

Shaded rows highlight CC composite score correlations.
** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).
[1]Signs of the WM Updating composite were reversed to match the directionality of the shifting and inhibition composites.

4.2.4.1 Shifting: Critical Region

There was no significant two-way interaction between violation condition and verb position, nor a three-way interaction between violation condition, verb position, and shifting composite score.

4.2.4.2 Shifting: Post-critical Region

There were no interaction effects present.

4.2.4.3 Inhibiting: Critical Region

The maximally converging model had the random slopes of violation condition by subject and item. There was a significant two-way interaction between the weaker violation contrast and verb position, beta estimate = 4.057×10^{-2}, $p = 0.016$ (Figure 9).

4.2.4.4 Inhibiting: Post-critical Region

There were no interaction effects present.

4.2.4.5 WM Updating: Critical Region

The maximally converging model had random slopes of violation condition by subject and item. A significant three-way interaction effect between the stronger violation contrast, verb position, and WM updating composite was found, beta estimate = 7.869×10^{-2}, $p < 0.000$. Better performance on the WM updating tasks was associated with greater adaptation to stronger violations during reading. Figure 10 displays the three-way interaction. Individuals with better performance on WM updating tasks showed longer reading times upon their first encounter with a stronger violation in the narratives, whereas individuals with poorer performance on the WM updating tasks did not show this pattern, and instead had longer reading times for the no violation condition upon first encounter.

4.2.4.6 WM Updating: Post-critical Region

The maximally converging model had random slopes of violation condition by subject and item. There was a significant three-way interaction effect, beta estimate = 6.69×10^{-2}, $p = 0.016$. The pattern of adaptation was similar to that in the critical region; there was a positive correlation between better performance on WM updating tasks and adaptation to stronger violations (see Figure 11).

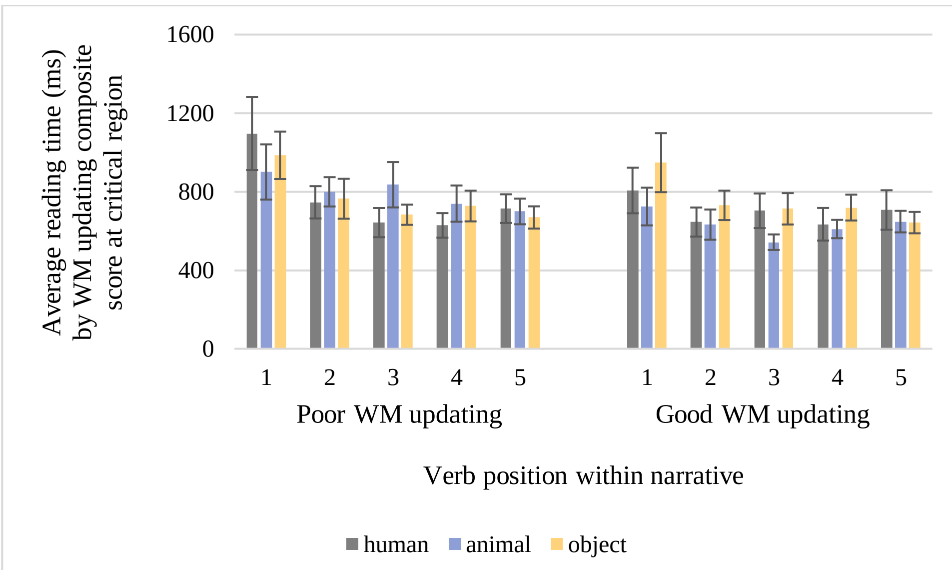


Figure 10 Mean reading times per violation condition at critical region graphed by median split on WM updating composite score. Error bars represent standard error.

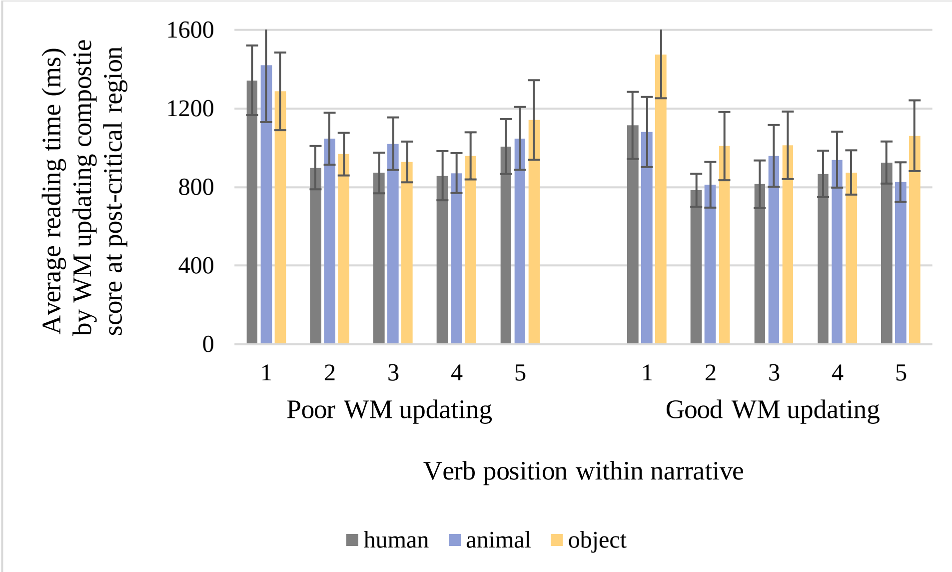


Figure 11 Mean reading times (ms) per violation condition at post-critical region graphed by median split on WM updating composite score. Error bars represent standard error.

4.2.4.7 Common CC: Critical Region

The maximally converging model had random slopes of violation condition by subject and item. A significant interaction between stronger violation contrast and verb position was found, beta estimate = -4.218×10^{-2}, p = 0.028. This appears to be a spurious interaction, as average means across the stronger violation condition and no violation condition at each verb position did not seem to differ (see Figure 12 for visual representation of the pattern). There was also a significant interaction between weaker violation contrast and verb position, beta estimate = 4.227×10^{-2}, p = 0.032. There was a marginal three-way interaction between stronger violation contrast, verb position, and common CC composite, estimated beta = 8.833×10^{-2}, p = 0.085. There was no three-way interaction effect showing a relationship between adaptation to weaker violations and WM updating ability.

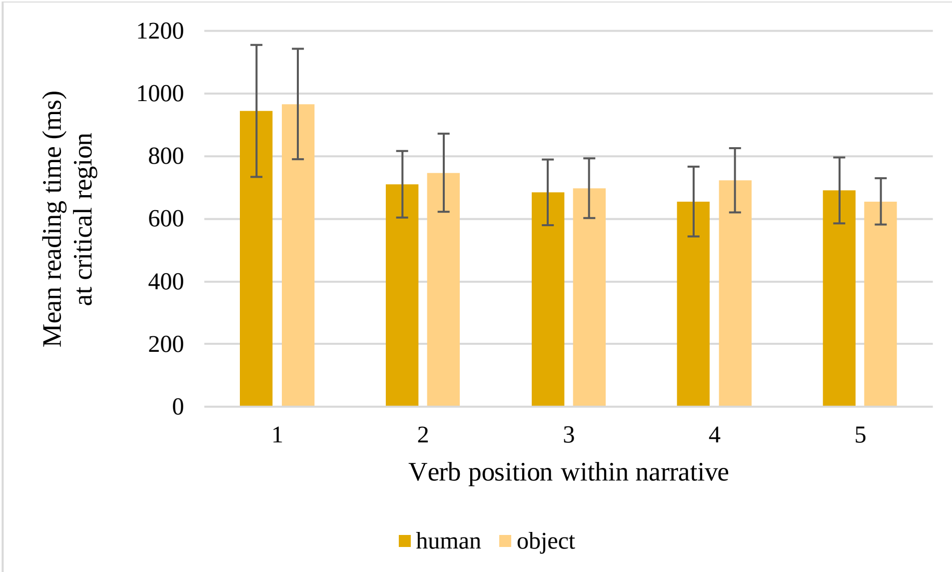


Figure 12 Mean reading times (ms) at the critical region for narratives with no violations (humans) and narratives with stronger violations (objects). Error bars represent standard error.

4.2.4.8 Common CC: Post-critical Region

There was no two-way or three-way interaction effects.

4.3 Discussion

Experiment 3 measured reading times for critical regions in narratives to investigate whether there was a relationship between domain-general cognitive control abilities and the magnitude of adaptation to different strengths of semantic violations. Because data collection was cut short due to the 2020 Covid-19 pandemic, there was a resulting small sample of individuals (N=12) for Experiment 3. This small sample size does not afford enough power to detect any true effects in the individual-differences analyses; hence, discussion of the results is tentative and general with respect to cognitive control, semantic adaptation, and their potential relationship, and no strong conclusions are drawn.

Adaptation within a narrative was examined through reading times on the critical regions, with decreased reading times with increased encounters with the violations indicating adaptation. Participant behavioral reading time patterns to stronger and weaker violations did not indicate adaptation within the narratives. Average reading time in the critical region for the stronger violation narrative condition was similar to average reading time for the no violation narrative condition, and average critical reading time for the weaker violation narrative condition was significantly *shorter* than the other two conditions. It is not clear why processing verbs indicating the first action of animal agents in a narrative would be faster than that of actions of human agents or inanimate object agents. This pattern does not follow adaptation to stronger violations observed in Experiment 2. It is difficult to speculate whether the reason for this difference is due to healthy aging, consequences of a small sample size (e.g. unrepresentative sample), a combination of the two, or something else entirely.

Individual difference analyses tested the relationship between cognitive control and semantic adaptation via perspective switching. Associations between cognitive control abilities and adaptation were examined through separate models, each including a cognitive control composite score for the cognitive processes of shifting, inhibiting, WM updating, as well as collective score of these processes, common cognitive control (CC). Results revealed a potential association between processing of stronger violations across a single narrative and WM updating ability. It seems that individuals who had better performance on the *Cued Shifting* and *Keep Track* tasks displayed typical disruption to their first encounter with a strong violation and, by their next encounter, showed decreased disruption. Hence, individuals with high WM updating cognitive profiles adapted to stronger violations. This behavioral pattern is likely due to a ceiling effect on the WM updating tasks. The interaction effect is likely driven by the fact that individuals with poorer WM updating performance showed a different pattern by which they had unusually longer reading times upon their first encounter with the actions of human animate agents in the narratives (no violation condition). It is unknown why this behavioral pattern was found for these individuals, as it cannot be attributed to overall longer reading time.

Even in the small sample of older neurotypical individuals, there was a relatively wide range of performance on tasks measuring shifting ability and ability to inhibit. As mentioned, there was less variability in performance across WM updating tasks, with a clear ceiling effect for both *Cued Shifting* and *Keep Track* for about half of the individuals. Correlational analyses revealed two interesting findings. First, there was no significant relationship between tasks measuring the same cognitive control ability. Tasks in the study were selected with the intention and assumption that they measure the same underlying construct (e.g. ability to inhibit). A lack of a correlation between similar tasks in the present sample does not support this claim. It might be that additional individual performance data on each task is necessary to observe these relationships. The second finding revealed moderate negative relationships among some inhibition tasks and WM updating tasks. This relationship in the data is surprising, as it is thought that WM updating and inhibition are more closely related than those to shifting ability, as WM is highly interdependent with inhibitory control via the two systems supporting the function of the other (e.g. Diamond, 2013, Trude & Tokowicz, 2011).

In sum, not much can be taken away from results of Experiment 3 due to the too-small sample size for individual difference analyses. Further speculation about cognitive control and semantic adaptation is considered in the General Discussion.

5.0 General Discussion

5.1 Adaptation Within Narratives

Whereas research has demonstrated our broad ability to adapt during language processing (e.g. Bradlow & Bent, 2008; Maye, Aslin, & Tanenhaus, 2008, for adaptation to speaker accent; e.g. Fraundorf & Jaeger, 2016, for adaptation to sentence structure; e.g. Boudewyn, Blalock, Long, & Swaab, 2019; Nieuwland & van Berkum, 2006, for adaptation to fantasy worlds during listening comprehension, e.g. Foy & Gerrig, 2014; Warren et al., 2008, during reading comprehension), the mechanisms responsible for adaptation during comprehension remain underspecified. The studies presented in this dissertation investigated specific questions related to the nature of adaptation mechanisms during reading comprehension. In all studies, semantic adaptation was examined through perspective switching: a change from relying on a set of beliefs for a real-world perspective to adopting beliefs of a fantasy-world perspective. The first aim of the studies tested whether behavioral evidence for the timing of perspective switching during reading aligned with principles of implicit, error-based learning, such that 1) stronger violations result in larger error than weaker violations and 2) larger error translates into greater and faster adaptation. Findings from this dissertation only partially support error-based learning and leave this question open. Experiment 1 indicated comprehenders switched perspectives rather quickly across a single narrative when short narratives contained stronger violations, and results from Experiment 2 indicated comprehenders switched to a fantasy-world perspective at a fast rate after encountering stronger violations. However, across Experiment 1 and 2, there was no evidence supporting adaptation to weaker violations.

Error-driven learning might be the mechanism that carries out semantic adaptation during reading. Data in this dissertation cannot speak to a direct comparison of the timing of adaptation to violations of varying strength. Hence I speculate here how future results–if finding adaptation differences–could be interpreted. Stronger violations present in narratives (i.e. inanimate agents performing human-specific animate actions) would be more unexpected than weaker violations (i.e. animal agents performing human-specific animate actions), and this greater mismatch in expectation would produce a larger error signal, which would lead to greater and faster adaptation. The idea that varying strengths of semantic violations are processed differently has been clearly established in the literature (e.g. Bohan, Leuthold, Hijikata, & Sanford, 2012; Rayner et al., 2004, for gradient disruption to good/poor-fit semantic violations), and a pattern of greater processing disruption to stronger than weaker violations was also observed in the current studies. Further, norming data in Experiment 1 confirmed the studies' violation manipulation by showing that comprehenders perceived stronger violations as being more unnatural than weaker violations in the narratives. The dissertation's novel contribution is its discussion regarding *how* current processing of semantic violations is meaningful for future processing of upcoming material. This work begins to test the question of whether an error-based explanation can account for the *how*; if the magnitude of the error signal during current processing drives the amount of adjustment (i.e. adaptation to a fantasy-world perspective) in subsequent comprehension in the narratives, with a stronger signal triggering more adjustment.

This line of reasoning partially supports error-driven adaptation in the dynamic generative framework of event comprehension (Kuperberg, 2016). Within the framework, the amount of prediction error drives the level of adaptation. If processing of a stronger violation in the narratives produces a large prediction error signal, then comprehenders would update their beliefs regarding the narrative following one encounter with a stronger violation. Belief updating in this case would be a large shift in beliefs because a stronger violation cue is informative that a new set of beliefs regarding events of the text would better capture meaning. According to the framework, a weaker violation cue would produce a smaller prediction error signal, and this signal would still be sent up through the model to inform future processing. But in this case, a weaker textual cue would not be sufficient to trigger an update in beliefs indicative of a perspective switch. Again though, it is important to be cautious with this interpretation and its application to the particular set of studies since there was no observed disruption to the weaker violations, suggesting an absence of prediction error.

5.2 Belief Updating vs. Implicit Learning

If error-driven learning is indeed responsible for semantic adaptation during reading, a question that this prompts is to what extent does adaptation lead to a complete change in expectations for upcoming text, such that previous expectations are ill-favored during processing. Thus far, I have discussed belief updating (e.g. Kuperberg, 2016) and implicit, error-based learning (e.g. Chang et al., 2012) accounts concerning adaptation as more or less similar, but a slight albeit significant difference between such models can have important implications for this question. While both accounts posit error as the driver of adaptation, they diverge in their explanation for how error adjusts expectations. The dynamic generative framework hypothesizes that error drives belief updating, such that a shift occurs from previously held beliefs to a new set of beliefs. This implies any previous expectations generated from prior beliefs are replaced with new expectations aligning with the new beliefs, and importantly, this shift reflects a change in frequency distribution.

Put another way, an increase in the probabilities of new expectations requires a decrease in probabilities of previous expectations. The process of belief updating in the dynamic generative framework could suggest a more complete form of adaptation during reading, and if so, would predict expectation violation effects following adaptation to previously held semantic expectations, such as the disruption effect observed in Nieuwland and van Berkum (2006, exp. 2) for real-world animacy expectations.

On the other hand, an error-based implicit learning account of adaptation hypothesizes that new expectations are strengthened with repeated exposure to error. A shift in weightings (i.e. probabilities) of expectations in these models does not necessarily suggest a strong, inverse relationship with weightings for other expectations to the degree of that of probabilities in belief updating accounts. Hence, an implicit learning model may support a form of adaptation that leads to shifted, but not absolute, new expectations. In this sense, there would be clear facilitative effects to the new expectations since these expectations are favored, but expectation violation effects for previous expectations during reading would not be predicted (for further discussion, see Dempsey, Liu, & Christianson, 2020). Evidence for implicit learning has been found in syntactic production (e.g. Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013) and syntactic comprehension (e.g. Fine & Jaeger, 2013; Fine, Jaeger, Farmer, & Qian, 2013) priming studies, and recently Dempsey and colleagues (2020) have provided strong evidence for a lack of expectation violation effects during syntactic comprehension. In this dissertation, study designs for Experiments 1 and 2 are unable to differentiate evidence for these forms of adaptation (i.e. do not test for expectation violation effects), but findings from Nieuwland and van Berkum (2006) suggests similar perspective switching to fantasy-world beliefs follows a more complete shift in expectations. This makes sense since a real-world perspective and fantasy-world perspective carry quite different statistical probabilities of events, though not mutually exclusive. Examining whether other scenarios of adaptation to semantic information reveal similar complete shifts in expectations may be beneficial in teasing apart these theories and the types of memory representations over which they operate.

A final point to consider regarding adaptation within fantasy narratives is how perspective switching is carried out over representations. As mentioned earlier, the dynamic generative framework mainly describes adaptation at a computational level of analysis, and only begins to explain adaptation mechanistically through its hypothesis of semantic feature (mis)matches for prediction error. This leaves open ambiguity regarding how updating of beliefs during a perspective switch involves a shift in activation and use of real-world knowledge representations in long-term memory to semantic representations for fantasy text. While representations for fantasy text cannot be completely disentangled from real-world knowledge representations (Cook & Guéraud, 2005, for discussion) since some aspects of general world knowledge are necessary to provide general structure to any story (e.g. temporal flow, character relations, etc.), a complete shift in expectations during belief updating implies that there is a tradeoff with the probabilities–and potentially the amount of activation–of each information source. The RI-Val model of comprehension (Cook & O'Brien, 2014) may inform this tradeoff with its explanation of competition between general world knowledge and contextual information for activation during initial processing. Competition between these sources allows for more of a continuum of activation for representations of each knowledge type, which consequently, may capture how a build-up of statistical information across textual cues (e.g. weaker violation cues) can lead to greater informativity for certain information and trigger adaptation.

5.3 Lack of Adaptation Across Narratives

While semantic adaptation to stronger violations was found within a narrative, across Experiments 1 and 2, there were mixed results whether adaptation to stronger violations occurred and no evidence for adaptation to weaker violation across narratives. This suggests that there may be a limit to the type (i.e. level) of cumulative, statistical information generated by comprehenders during processing. Behavioral evidence for a lack of adaptation across narratives in the studies could mean that comprehenders did not extend their beliefs about a single narrative to other narratives that contained similar cues (i.e. beliefs regarding one inanimate object agent carrying out human actions to all narratives containing object agents). Differing results across Experiments 1 and 2 could instead reflect task demands, which could have led participants to employ strategies in Experiment 1 that simulated adaptive behavior. Specifically, the cumulative cloze task had participants produce their best guess response for the next word in the narrative. This conscious and continual action was always followed with presentation of a human-specific verb in the narratives, regardless of the type of main character of the story. Because of this, and in conjunction with low-probability events described in the narratives, participants may have incidentally learned a pattern to simply expect to read *odd* things across the experiment. Thus, instead of adapting to the semantic violations present in the narratives, participants may have applied a general strategy to produce odd actions.

Different probabilities for encountering fantasy-related actions within a given narrative and across narratives in Experiments 1 and 2 could have also contributed to the findings. The likelihood of an inanimate object or animal agent performing additional fantasy-related actions was 100% within a narrative, whereas there was a lower likelihood that participants would continue to encounter fantasy-related information across narratives due to a portion of the narratives describing a human agent. Thus, fantasy-related predictions within a story would have been more successful than initial fantasy-related predictions in subsequent narratives. If this was true, and assuming comprehension processes follow higher-level statistical information (e.g. probabilities between narratives), then perspective switching may only occur after a certain likelihood threshold is reached.

5.4 Cognitive Control and Adaptation

The final aim of this dissertation was to investigate the cognitive underpinnings associated with semantic adaptation mechanisms. If comprehenders adapt their processing by tracking informativity changes of textual cues present in the reading environment, then a certain level of monitoring and control may be required to recognize when and to manage how a change in information use should occur. Such control may include adopting new probabilities, ignoring unhelpful information, and revising beliefs according to what is informative. Experiment 3 related individual participant performance on tests of domain-general cognitive control ability (shifting, inhibiting, and WM updating) to their adaptation during a self-paced reading task. Findings from Experiment 3 are not elaborated upon here due to its too-small sample size to draw any telling conclusions, and instead, cognitive control and various forms of adaptation during language processing are discussed more broadly.

Two simple tasks were chosen based on the literature (Allen et al., 2012; Simic, 2019) in Experiment 3 to measure each of the three core cognitive control processes of shifting, inhibition, and WM updating (Miyake et al., 2000). These tasks are thought to primarily tap each control process separately, though with the nature of cognitive control being to act upon (i.e. control) other cognitive processes, pure measures of cognitive control are difficult to obtain (Simic, 2019). To best mitigate this task impurity problem, composite scores were created for each separable cognitive control process as well as an aggregated, common cognitive control score (Miyake et al., 2000). This methodological process is important and necessary to implement with cognitive control data in order to better understand the extent by which these core processes share common components and are interrelated (i.e. common factor) and the degree that they involve specific, distinguishable mechanisms beyond those shared (i.e. specific factor, for further discussion on unity/diversity framework for cognitive control, see Miyake et al., 2000; Miyake & Friedman, 2012).

Continued research examining the potential role of cognitive control during language processing can reveal whether various forms of adaptation require none, similar, or differing engagement across the core processes of shifting, inhibiting, and WM updating. And this information could yield useful in targeting best-practice approaches for both improvement in healthy (e.g. Novick et al., 2014 for cognitive control training on garden-path recovery) and impaired (e.g. Simic et al., 2019 for cognitive control as a predictor of language recovery post-stroke) language processing. Much of the research on cognitive control during neurotypical language comprehension has focused on working memory control over adaptation during syntactic processing (e.g. Novick et al., 2010): adaptation via conflict resolution resulting from misinterpretation of sentence structure (e.g. garden-path effect). This form of adaptation is one of revision, by which comprehenders must change from one set of syntactic representations to another for successful comprehension. This conceptualization of syntactic adaptation during processing can encompass other forms of change as well beyond that of semantic adaptation as discussed in this dissertation. Another example of such change is switching from one language to another (i.e.

bilingualism). In fact, domain-general cognitive abilities have been implicated in second-language learning. For example, Trude and Tokowicz (2011) found in a production task that individuals with higher working memory showed improved ability to inhibit their first language when producing words in a second language. There is still much to be understood about the contributions of domain-general cognitive control to various aspects of language processing, particularly the mechanisms responsible for adaptation. Experiment 3 of this dissertation can be considered a step in this direction. It is worth mentioning while there is evidence supporting a strong role of domain-general cognitive control in language processing and particularly in predictive processing, as mentioned in this dissertation, there is also evidence suggesting that many of the predictive and violating effects observed in psycholinguistic research could be attributed to language-specific mechanisms (see Ryskin et al., 2020 for review; Shain et al., 2020 for example).

5.5 Strengths and Limitations

The studies in this dissertation investigated semantic adaptation through perspective switching. This experimental decision provided a direct way to measure adaptation, as there was a clear point during processing to behaviorally examine when comprehenders moved from relying on a real-world perspective to relying on a fantasy-world perspective. Another strength of this dissertation lies in its purposeful connection of distinct literatures to advance research. Background knowledge for the dissertation combined the fantasy text processing literature–often considered a niche literature–with ideas from the computational, error-based learning literature. This combination of ideas afforded testing of specific, theory-driven (i.e. dynamic generative framework of event comprehension; Kuperberg, 2016) hypotheses regarding adaptation during reading. A third strength comes from the use of multiple approaches to investigate semantic adaptation, including both a language-specific and more general cognitive approach. The dissertation tested properties of language-specific adaptation mechanisms of comprehension as well as tested the type(s) of domain-general control over these mechanisms. Experiment 3 is, to my knowledge, one of the first studies to look for a relationship between semantic adaptation and cognitive control.

The primary limitation of the dissertation is the small sample size for Experiment 3. Future replication in a larger group of older participants would be required to determine whether there is a relationship (and the nature of that relationship) between participant cognitive control profiles and their pattern of observed semantic adaptation during reading. In addition, the studies in this dissertation varied in their analyses for evidence of semantic adaptation (i.e. across all verb positions, only across first two verb encounters), which could have influenced conclusions drawn regarding error-driven learning and adaptation. Finally, data from the present set of studies only addresses broad questions regarding error-based learning; future research investigating semantic adaptation is needed to test hypotheses that would differentiate between implicit learning and belief updating theories.

5.6 Conclusion

Findings from this dissertation contribute to research which aims to understand the mechanisms driving semantic adaptation during reading. Results from Experiments 1 and 2 suggest that error-based learning might be responsible for adaptation to a fantasy-world perspective, but this suggestion is limited due to a lack of observed adaptation to weaker violations in the narratives. Experiment 3 findings are preliminary but hint that domain-general cognitive control processes might underlie our ability to perspective switch, with WM updating potentially playing an important role in the updating of beliefs. A better understanding of how adaptation is accomplished by comprehension processes is necessary for any complete theory of language comprehension. Current theories are beginning to capture the complex interplay of different sources of information and their influence on processing, and future research on adaptation, specifically the time-course by which comprehenders change their reliance on information, can aide in this understanding.

Appendix A Experimental Stimuli for Experiments 1-3

Note: The set of stimuli presented here reflect stimuli used in Experiment 2. Slight modifications were made to the set of stimuli between Experiments 1 and 2, and Experiments 2 and 3 to accommodate presentation program displays (see descriptions in text).

1

On a chilly morning, a **fisherman/worm/hook** was sipping a latte in a coffee shop.
A fellow customer kept looking his way.
The **fisherman/worm/hook** smiled at the customer and blushed bright pink.
The customer was seated at a far table, so the **fisherman/worm/hook** walked toward the customer.
The **fisherman/worm/hook** mumbled, "How is your cup of joe?"

2

A few weeks ago, a **yodeler/goat/mountain** was interviewed by the local news.
There had been a horrible car accident a few days before.
With the camera rolling, the **yodeler/goat/mountain** spoke about the quick EMT response and he explained that the rescuers saved many lives.
The **yodeler/goat/mountain** answered all the reporter's questions as truthfully as possible.
The **yodeler/goat/mountain** was also polite to the reporter and did well in the interview.

3

There once was a **magician/rabbit/hat** that owned a new fancy restaurant downtown.
On opening night, the **magician/rabbit/hat** invited all his friends, and he advertised his dishes to local food critics.
The **magician/rabbit/hat** was very thrilled, and the night went smoothly.
The **magician/rabbit/hat** declared, *Maybe this was a success*, as the last customer left the establishment.
This was the first time that the **magician/rabbit/hat** ventured into the restaurant business.

4

On an ordinary Monday, a **beekeeper/bee/honeycomb** strolled into the bank.
The **beekeeper/bee/honeycomb** requested a withdrawal slip so she could take out $100.
The bank teller asked for the account information.
The **beekeeper/bee/honeycomb** rummaged for her wallet and whispered her account number to the teller.
The **beekeeper/bee/honeycomb** was buying a birthday gift for her daughter.

5

After a bout of sickness, a **chef/lobster/butter stick** finally scheduled a doctor's appointment.
Luckily, an appointment was available the very next day.
The **chef/lobster/butter stick** had sneezed way too many times the past week.
After the examination, the **chef/lobster/butter stick** was prescribed medication and he listened to the doctor's advice about getting lots of rest.
For the next week, though, the **chef/lobster/butter stick** blew his runny nose continually.

6

Across the world, there was a **diver/shark/cage** that announced all the baseball games in his city.
At the start of a big match-up, the **diver/shark/cage** introduced himself and his partner to the crowd and then sung their national anthem.
After the anthem, the **diver/shark/cage** discussed all the players on each team and their statistics.
It was the fiftieth game of the season.
The weather became ugly, and the **diver/shark/cage** broadcasted a delay of game after the third inning.

7

This summer, a **trainer/dolphin/whistle** <u>volunteered</u> as a camp counselor in the Rockies.

The **trainer/dolphin/whistle** <u>remembered</u> her favorite activities as a camper at the same summer camp many years earlier.

These memories included the young **trainer/dolphin/whistle** <u>swinging</u> on the tire and <u>canoeing</u> on the lake.

The campers always had a lot of fun with their counselors at this camp.

The **trainer/dolphin/whistle** <u>prayed</u> that no camper would get injured during the week.

8

One day in a studio in California, a **policeman/German shepherd/gun** was the <u>host</u> of a daytime television game show.

They were currently filming a holiday episode.

As the contestant played the first game, the **policeman/German shepherd/gun** <u>gestured</u> back to the director for a commercial break.

The **policeman/German shepherd/gun** <u>calculated</u> the time until they were back on the air and then he <u>commanded</u> the audio engineer to turn down the Christmas music.

As fake snow fell on the stage, the **policeman/German shepherd/gun** <u>yelled</u>, “And we’re back!”

9

In a far country, an/a **Eskimo/dog/sled** <u>taught</u> geology at a school on a hilltop.

To make his class fun, the **Eskimo/dog/sled** <u>buil</u>t a volcano and <u>poured</u> baking soda and vinegar into it.

The **Eskimo/dog/sled** also <u>ordered</u> geodes and <u>demonstrated</u> how to break them.

The students enjoyed these lessons very much.

And the young children were inspired to pursue science as a career.

10

Last Thursday, a **farmer/cow/milk bottle** <u>exercised</u> at the gym down the street.

The **farmer/cow/milk bottle** <u>deadlifted</u> 100 pounds and then <u>jogged</u> a couple miles on the treadmill.

The leg press was another planned workout for the day.

Afterwards, the **farmer/cow/milk bottle** <u>wiped</u> down all the machines**.**

The **farmer/cow/milk bottle** <u>crawled</u> out of the gym exhausted.

11

Some time ago, a **zookeeper/lion/tree** <u>bought</u> tickets to an amusement park.

The trip had been scheduled for months.

On every spinning ride, the **zookeeper/lion/tree** <u>giggled</u> uncontrollably and <u>announced</u> how dizzy he had become.

The **zookeeper/lion/tree** <u>reminisced</u> about how different the park looked only a couple years ago.

The **zookeeper/lion/tree** was <u>impressed</u> with all the updated rides and facilities.

12

At one point, a **jockey/horse/saddle** went <u>shopping</u> at a large grocery store.

The **jockey/horse/saddle** was <u>disappointed</u> to find the store overly crowded.

The **jockey/horse/saddle** <u>muttered</u> some not-so-nice words under his breath.

After a couple seconds of contemplation, the **jockey/horse/saddle** <u>decided</u> that he would try another store.

The **jockey/horse/saddle** <u>drove</u> out of the parking lot in a hurry.

13

On a highly-anticipated movie release night, a **hunter/deer/gun** <u>drove</u> to the theater.

The **hunter/deer/gun** <u>asked</u> someone in line if there were any tickets available for the movie.

That movie was sold out, so the **hunter/deer/gun** <u>chose</u> another action film and then <u>handed</u> the worker his ticket to scan.

But inside the theater, an alternative presented itself.

The **hunter/deer/gun** <u>slithered</u> into the wrong theater room to watch the newest movie.

14

In the olden days, a **farmer/chicken/feed barrel** was employed as the town's pharmacist on Saturdays.
Most of the time, the **farmer/chicken/feed barrel** counted medication and talked with his neighbors.
Even though the hours were long, it was a good job.
The **farmer/chicken/feed barrel** was notorious for gossiping about the town and bragging about his farm any chance he could get.
But everyone in town was happy to have such a reliable pharmacist.

15

On Friday, a **cowboy/horse/saddle** consulted with his therapist.
The therapist was a specialist in treating mood swings.
After the session, the **cowboy/horse/saddle** questioned his therapist about things he could do to control his mood.
The **cowboy/horse/saddle** laughed when his therapist suggested more exercise and he cringed when the therapist said to avoid alcohol.
The **cowboy/horse/saddle** begrudgingly reflected, *there goes my Sunday football routine*.

16

One time a **veterinarian/dog/stethoscope** was gossiping about a telemarketer who frequently called the house.
When the phone rang again, the **veterinarian/dog/stethoscope** brainstormed what kind of joke she would play on the telemarketer.
The **veterinarian/dog/stethoscope** pretended to be a parrot and she repeated every word that the telemarketer spoke.
Then the **veterinarian/dog/stethoscope** advised the telemarketer to leave a message after the beep.
Confused, the telemarketer hung up.

17

There once was a **man/cat/tree** who assembled tents for the traveling circus.
It was a tough job but the tent workers could watch the circus for free.
Though all aspects of the circus were great, the **man/cat/tree** whistled rather boisterously and clapped the loudest for the elephants' performance.
When the act ended, the **man/cat/tree** cheered for an encore.
The **man/cat/tree** envisioned an even bigger circus next year.

18

For the first time, a **dog show handler/dog/trophy** borrowed a book from the library.
The book was quite long but contained important information.
The **dog show handler/dog/trophy** read the entire book in two hours.
The **dog show handler/dog/trophy** wrote a lot of notes and she googled for more books to read.
The next day, the **dog show handler/dog/trophy** ordered an Uber to go back to the library.

19

A **marine biologist/fish/boat** jogged into a department store to start work.
The next shift started in five minutes.
Once her shift started, the **marine biologist/fish/boat** cataloged new inventory that was in boxes.
The **marine biologist/fish/boat** ironed the clothes and requested that she be put on another project.
The manager shook his head, for the **marine biologist/fish/boat** had forgotten to unpack a few boxes.

20

After an interesting morning, a **blind man/service dog/walking stick** telephoned his museum curator friend**.**
The curator had been a close friend since high school.
The **blind man/service dog/walking stick** had recently purchased a piece of artwork and inquired if his friend would come over to examine it.

After what seemed like hours, the **blind man/service dog/walking stick** <u>texted</u> the curator for his whereabouts. Then the **blind man/service dog/walking stick** <u>mopped</u> the kitchen floor while he waited for his friend.

Appendix B Summary of Mixed-Effects Models for Experiment 1

Adaptation to Stronger Violations

Within-Narrative Adaptation

Model 1: Stronger Violation Response ~ 1 + Violation Condition*Verb Position +
(1 + Violation Condition + Verb Position | Subject) + (1 + V. Condition | Item)

Predictor	Coefficient	SE	Z	p
Intercept	0.492	0.213	2.314	0.021
Violation (V.) Condition	-1.639	0.294	-5.570	< 0.001
Verb (V.) Position	0.259	0.065	3.979	< 0.001
V. Condition*V. Position	0.447	0.088	5.089	< 0.001

Across-Narrative Adaptation

Model 2: Stronger Violation Response ~ 1 + Violation Condition*Narrative Position + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-0.278	0.393	-0.712	NS
V. Condition	-2.600	0.666	-3.906	< 0.001
Narrative (N.) Position	0.082	0.029	2.841	0.005
V. Condition*N. Position	0.118	0.057	2.052	0.040

Adaptation to Weaker Violations

Within-Narrative Adaptation

Model 3: Weaker Violation Response ~ 1 + Violation Condition*Verb Position +
(1 + Violation Condition + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-1.219	0.178	-6.849	< 0.001
V. Condition	0.788	0.288	2.740	0.006
V. Position	0.159	0.041	3.900	< 0.001
V. Condition*V. Position	-0.160	0.092	-1.736	0.082

Across-Narrative Adaptation

Model 4: Weaker Violation Response ~ 1 + Violation Condition*Narrative Position + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-2.179	0.375	-5.810	< 0.001
V. Condition	1.688	0.756	2.234	0.026
N. Position	0.034	0.030	1.160	NS
V. Condition*N. Position	-0.025	0.062	-0.401	NS

Appendix C Filler Stimuli for Experiments 2, 3

1f

One fall day, a barista was hiking with his friend.
The friend was whistling loudly in the woods.
The barista rolled his eyes at his friend and then almost tripped over a hidden tree root.
Then, he laughed nervously so that his friend glanced at him with a quizzical look.
The barista muttered, “The tree almost got me.”

2f

One afternoon a radio host was picnicking at Yellowstone.
The hot springs were very beautiful and it wasn’t too cold.
The radio host draped her blanket on the warm ground and arranged her food containers in a neat row.
The radio host photographed herself in front of the hot spring.
Then, she casually prepared a plate and ate her meal in peace.

3f

One evening a hotel manager attended a musical.
The hotel manager gasped when she heard the opener and she grinned at the singer.
She knew the song and liked it very much.
But the hotel manager was apathetic to the rest of the performance.
The musical as a whole left the hotel manager confused as she strolled back to the hotel.

4f

On Wednesday, an astronaut was wandering around a construction site.
The astronaut was scanning the area for a ring he lost the day before.
He was startled when a construction worker yelled, “Sir, this is a dangerous area.”
The astronaut shrugged his shoulders and replied back, “I need to find my ring.”
The construction worker quickly hid his hand behind his back.

5f

While on his lunch break, a traffic cop strolled into a shoe store.
He was really hoping they had some designer loafers for an upcoming formal event.
He flagged down the attendant to ask if they had any promotional discounts.
She explained that the discounts only applied to teen shoes and the cop nodded his understanding.
The traffic cop then typed a quick message to his partner to please buy him the loafers online

6f

After a long Monday, a painter was relaxing in the local firehall.
She read a sign stating that bingo was happening that night.
The painter judged her odds of winning and concluded that she would stay for the game.
She had not played since his high school days.
The painter was eager to see the winning prize.

7f

One afternoon a black-belt was piloting his hot air balloon.
The black-belt steered his hot air balloon toward a barley field.
Gazing over the edge, he tallied the number of sheep in his landing path.
“Drat,” he sputtered, “why do I always forget about the sheep?”
He navigated the hot air balloon away from the herd as best he could.

8f

Once, a mechanical engineer was singing karaoke in a bar.
She was amazed at all the song options she could choose from on the machine.
The engineer danced energetically while she sang and her friends joined in the fun too.
The mechanical engineer motioned to her one friend to begin the grand finale.
The engineer and her friend hurdled the bar stools and strummed their air guitars until the music stopped.

9f

Many days ago a roller coaster enthusiast ordered a steak at a steakhouse.
She was dining alone and had just journaled about her experience on the new wooden coaster at her nearby theme park.
Upon reflection of the ride, the roller coaster enthusiast determined that she was no longer hungry for a steak.
When the waitress returned, the enthusiast pointed to the menu.
“Can I change my order and just get a milkshake instead?”

10f

In the city, a CEO visited his favorite bakery.
The CEO waved to the baker and commented, “We’re out of wheat bread again at home.”
The baker shifted his weight uncomfortably and said, “So are we.”
The CEO cursed under his breath.
He gestured toward the shelf in resignation, responding, “I guess I’ll take sourdough.”

Appendix D Summary of Initial Mixed-Effects Models for Experiment 2

Note: The initial models included the full range for the fixed effect of verb position (verbs 1-5 in a narrative).

Within-Narrative Adaptation

Gaze duration

Critical region

Model 1: Gaze duration ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	307.144	9.041	64.155	33.971	< 0.001
Violation (V.) Condition	-4.850	5.185	17.917	-0.936	NS
Experiment	-19.037	14.612	66.439	-1.303	NS
Verb (V.) Position	-19.435	1.783	236.618	-10.903	< 0.001
V. Condition*Experiment	-12.918	9.398	5417.877	-1.375	NS
V. Condition*V. Position	-4.068	3.318	5417.755	-1.266	NS
Experiment*V. Position	-2.380	3.561	235.803	-0.668	NS
V. Condition*Exp.*V. Position	1.438	6.636	5405.333	0.217	NS

Post-critical region

Model 2: Gaze duration ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	412.610	14.458	55.404	28.539	< 0.001
V. Condition	-4.415	7.065	56.328	-0.625	NS
Experiment	1.561	21.273	67.429	0.073	NS
V. Position	-12.790	2.432	62.221	-5.259	< 0.001
V. Condition*Experiment	-7.968	13.023	5688.586	-0.612	NS
V. Condition*V. Position	1.126	4.567	5694.521	0.247	NS
Experiment*V. Position	-7.128	4.862	62.154	-1.466	NS
V. Condition*Exp.*V. Position	8.244	9.137	5695.507	0.902	NS

First-pass regressions out

Critical region

Model 3: First-pass regression out ~ 1 + Violation Condition*Experiment*Verb Position (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	*Z*	p
Intercept	-1.341	0.093	-14.413	< 0.001
V. Condition	0.340	0.074	4.573	< 0.001
Experiment	-0.036	0.154	-0.231	NS
V. Position	0.053	0.028	1.903	0.057
V. Condition*Experiment	-0.330	0.136	-2.430	0.015
V. Condition*V. Position	-0.031	0.048	-0.642	NS
Experiment*V. Position	0.160	0.055	2.910	0.004
V. Condition*Exp.*V. Position	-0.126	0.096	-1.313	0.189

Post-critical region

Model 4: First-pass regressions out ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 + Verb Position | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-1.256	0.101	-12.378	< 0.001
V. Condition	0.244	0.064	3.808	< 0.001
Experiment	-0.233	0.132	-1.758	0.079
V. Position	0.005	0.023	0.232	NS
V. Condition*Experiment	-0.358	0.128	-2.789	0.005
V. Condition*V. Position	-0.046	0.045	-1.026	NS
Experiment*V. Position	0.150	0.045	3.336	< 0.001
V. Condition*Experiment*. Position	-0.128	0.090	-1.418	NS

Go-past

Critical region

Model 5: Go-past time ~ 1 + Violation Condition*Experiment*Verb Position (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	t	p
Intercept	481.495	21.622	63.351	22.269	< 0.001
V. Condition	35.832	15.794	18.259	2.269	0.036
Experiment	-44.921	38.485	61.896	-1.167	NS
V. Position	-16.707	5.449	312.262	-3.066	0.002
V. Condition*Experiment	-41.402	29.556	5421.485	-1.401	NS
V. Condition*V. Position	-13.786	10.441	5416.016	-1.320	NS
Experiment*V. Position	22.253	10.889	311.219	2.044	0.042
V. Condition*Experiment*V. Position	-6.747	20.881	5402.041	-0.323	NS

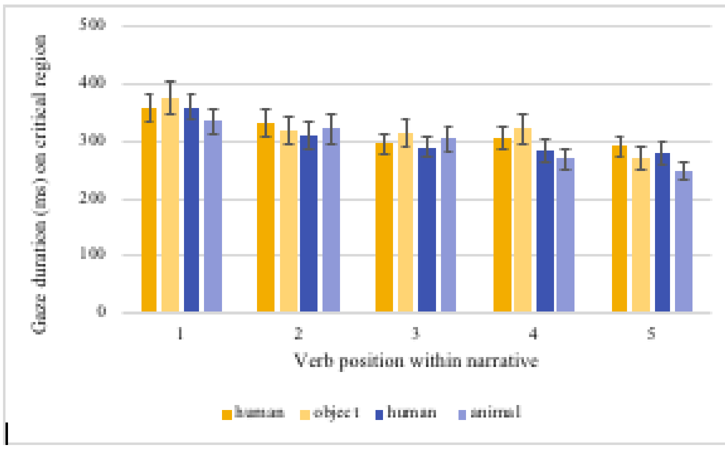
Post-critical region

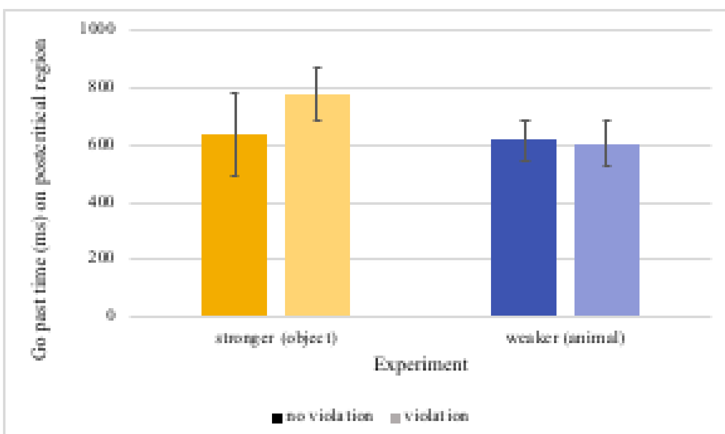
Model 6: Go-past time ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

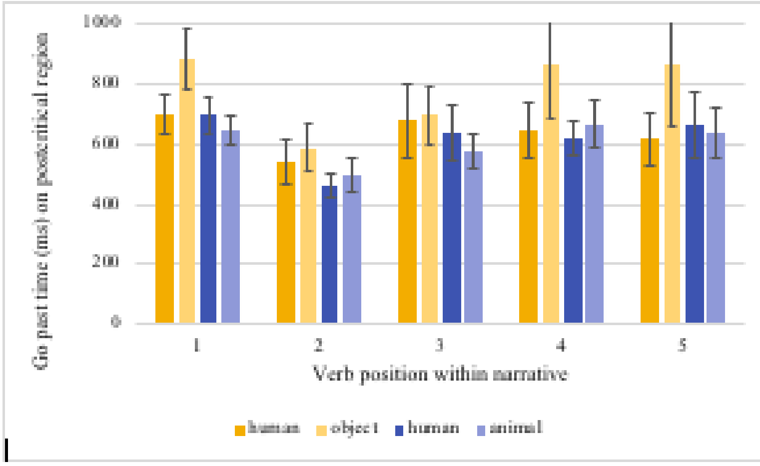
Predictor	Coefficient	SE	df	t	p
Intercept	665.201	34.054	43.101	19.534	< 0.001
V. Condition	66.589	21.657	1671.244	3.075	0.002
Experiment	-100.650	47.873	64.440	-2.102	0.039
V. Position	10.541	7.897	322.450	1.335	NS
V. Condition*Experiment	-148.279	43.180	5742.037	-3.434	0.001
V. Condition*V. Position	17.538	15.144	5736.890	1.158	NS
Experiment*V. Position	6.646	15.789	322.093	0.421	NS
V. Condition*Experiment *V. Position	-26.162	30.295	5738.054	-0.864	NS

Appendix E Graphs from Initial Analyses for Experiment 2

Note: The select graphs from the initial analyses are presented here as comparison to Figures 4-6 to show the full pattern of data across all verb positions (1-5).







1000
800
600
400
200
0
1
2
3
4
5
Go past time (ms) on postcritical region
Verb position within narrative
human
object
human
animal

Appendix F Summary of Mixed-Effects Models for Experiment 2

Note: Subsequent models for within-narrative adaptation for Experiment 2 included only the 1st and 2nd verb encounters in a narrative for the fixed factor of verb position.

Within-Narrative Adaptation

Gaze duration

Critical region

Model 1: Gaze duration ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	333.679	14.522	35.331	22.978	< 0.001
Violation (V.) Condition	-6.085	8.807	16.228	-0.691	NS
Experiment	-13.130	17.792	64.420	-0.738	NS
Verb (V.) Position	37.049	9.938	68.468	3.728	< 0.001
V. Condition*Experiment	-14.233	15.828	2107.141	-0.899	NS
V. Condition*V. Position	1.425	15.810	2105.703	0.090	NS
Experiment*V. Position	-9.989	19.876	68.454	-0.503	NS
V. Condition*Experiment*V. Position	-68.175	31.631	2105.726	-2.155	0.031

Post-critical region

Model 2: Gaze duration ~ 1 + V. Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	422.695	23.077	29.853	18.317	< 0.001
V. Condition	-3.824	10.070	2239.241	-0.380	NS
Experiment	10.492	23.717	67.521	0.442	NS
V. Position	130.445	10.915	191.296	11.951	< 0.001
V. Condition*Experiment	-16.461	20.130	2239.241	-2.191	0.029
V. Condition*V. Position	-44.013	20.092	2238.173	-0.818	NS
Experiment*V. Position	36.752	21.828	191.215	1.684	0.094
V. Condition*Experiment*V. Position	29.249	40.186	2238.188	0.728	NS

First-pass regressions out

Critical region

Model 3: First-pass regression out ~ 1 + Violation Condition*Experiment*Verb Position (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	*Z*	p
Intercept	-1.443	0.135	-10.729	< 0.001
V. Condition	0.470	0.110	4.290	< 0.001
Experiment	-0.322	0.193	-1.666	0.096
V. Position	-0.235	0.109	-2.144	0.032
V. Condition*Experiment	-0.099	0.219	-0.450	NS
V. Condition*V. Position	0.050	0.219	0.228	NS
Experiment*V. Position	-0.498	0.219	-2.278	0.023
V. Condition*Experiment*V. Position	0.557	0.437	1.273	NS

Post-critical region

Model 4: First-pass regressions out ~ 1 + Condition*Experiment*Verb Position + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-1.375	0.123	-11.189	< 0.001
V. Condition	0.298	0.106	2.814	0.005
Experiment	-0.515	0.173	-2.983	0.003
V. Position	0.518	0.106	4.889	< 0.001
V. Condition*Experiment	-0.180	0.211	-0.852	NS
V. Condition*V. Position	0.359	0.211	1.698	0.090
Experiment*V. Position	-0.016	0.211	-0.076	NS
V. Condition*Experiment*V. Position	-0.450	0.422	-1.065	NS

Go-past

Critical region

Model 5: Go-past time ~ 1 + Violation Condition*Experiment*Verb Position (1 + Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	t	p
Intercept	501.532	24.159	43.591	20.759	< 0.001
V. Condition	73.373	29.040	18.881	2.527	0.021
Experiment	-58.678	36.658	53.714	-1.601	NS
Verb Position	-18.171	21.785	80.734	-0.834	NS
V. Condition*Experiment	-50.318	36.625	2141.620	-1.374	NS
V. Condition*V. Position	-36.013	36.612	2141.490	-0.984	NS
Experiment*V. Position	-54.778	43.575	80.764	-1.257	NS
V. Condition*Experiment*V. Position	1.625	73.241	2141.591	0.022	NS

Post-critical region

Model 6: Go-past time ~ 1 + Violation Condition*Experiment*Verb Position + (1 + Verb Position | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	t	p
Intercept	628.170	38.220	30.400	16.434	< 0.001
V. Condition	60.130	22.730	2174.450	2.646	0.008
Experiment	-109.780	43.110	60.370	-2.547	0.013
Verb Position	213.690	26.760	55.860	7.985	< 0.001
V. Condition*Experiment	-111.230	45.430	2175.280	-2.448	0.014
V. Condition*V. Position	37.100	45.350	271.150	0.818	NS
Experiment*V. Position	-45.080	53.520	55.860	-0.842	NS
V. Condition*Experiment*V. Position	-215.700	90.700	2171.310	-2.378	0.018

Note: Models for across-narrative adaptation for Experiment 2 included the first verb encounters in the first three and last three presented narratives per participant.

Across-Narrative Adaptation

Gaze duration

Critical region

Model 7: Gaze duration ~ 1 + Violation Condition*Experiment*Narrative Position + (1 + Narrative Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	358.981	19.631	27.716	18.286	< 0.001
V. Condition	-3.188	18.916	15.382	-0.169	NS
Experiment	-44.113	22.522	58.821	-1.959	0.055
Narrative (N.) Position	-19.516	15.844	61.858	-1.232	NS
V. Condition*Experiment	-62.789	28.066	554.395	-2.237	0.026
Experiment*N. Position	-36.716	28.270	561.555	-1.299	NS
V. Condition*N. Position	-12.962	31.651	61.234	-0.410	NS
V. Condition*Experiment*N. Position	36.556	56.783	563.049	0.644	NS

Post-critical region

Model 8: Gaze duration ~ 1 + Violation Condition*Experiment*Narrative Position + (1 + Narrative Position | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	477.309	36.859	22.977	12.950	< 0.001
V. Condition	-20.130	20.090	617.433	-1.002	NS
Experiment	20.406	30.701	68.024	0.665	NS
N. Position	-1.308	20.318	522.120	-0.064	NS
V. Condition*Experiment	24.392	40.133	618.150	0.608	NS
V. Condition*N. Position	-0.398	40.390	618.245	-0.010	NS
Experiment*N. Position	64.555	40.282	516.096	1.603	NS
V. Condition*Experiment*N. Position	104.921	81.246	619.808	1.291	NS

First-pass regressions out

Critical region

A minimal model for the critical region would not converge.

Post-critical region

Model 10: First-pass regressions out ~ 1 + Violation Condition*Experiment*Narrative Position + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	Z	p
Intercept	-1.123	0.166	-6.785	< 0.001
V. Condition	0.716	0.185	3.858	< 0.001
Experiment	-0.445	0.220	-2.027	0.043
N. Position	0.163	0.184	0.884	NS
V. Condition*Experiment	-0.219	0.367	-0.598	NS
V. Condition*N. Position	0.047	0.368	0.127	NS
Experiment*N. Position	-0.423	0.367	-1.154	NS
V. Condition*Experiment*N. Position	0.051	0.738	0.069	NS

Go-past

Critical region

Model 11: Go-past time ~ 1 + Violation Condition*Experiment*Narrative Position (1 + Narrative Position | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	477.780	24.200	28.450	19.741	< 0.001
V. Condition	49.840	25.430	564.180	1.960	0.051
Experiment	-95.220	34.690	69.420	-2.745	0.008
N. Position	19.000	27.150	66.370	0.700	NS
V. Condition*Experiment	-75.780	50.790	560.690	-1.492	NS
V. Condition*N. Position	10.930	51.060	567.990	0.214	NS
Experiment*N. Position	-81.260	54.120	65.360	-1.501	NS
V. Condition*Experiment*N. Position	74.590	102.630	572.09	0.727	NS

Post-critical region

Model 12: Go-past time ~ 1 + Violation Condition*Experiment*Narrative Position + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	725.200	62.590	22.670	11.587	< 0.001
V. Condition	115.460	40.3350	617.880	2.862	0.004
Experiment	-142.840	56.900	66.290	-2.510	0.015
N. Position	10.760	40.630	624.060	0.265	NS
V. Condition*Experiment	-160.970	80.610	618.670	-1.997	0.046
V. Condition*N. Position	45.990	81.110	619.140	0.567	NS
Experiment*N. Position	-10.190	80.560	621.480	-0.127	NS
V. Condition*Experiment*N. Position	178.610	163.110	621.170	1.095	NS

Appendix G Self-Paced Reading Parsing of Experiment 3 Stimuli

Note: Asterisks (*) indicate self-paced reading parsed segments. Underlined segments denote critical regions; italicized segments denote post-critical regions.

1

On a chilly morning,* a fisherman/worm/hook* was sipping* _a latte_* in a coffee shop.
A fellow customer* kept looking his way.
The fisherman/worm/hook* smiled* _at the customer and_* blushed* _bright pink_.
The customer* was seated far away,* so the fisherman/worm/hook* walked* _toward the customer_.
The fisherman/worm/hook* mumbled,* _"How is your cup of joe?"_

2

A few weeks ago,* a yodeler/goat/mountain* was interviewed* _by the local news_.
There had been* a horrible accident* a few days before.
The yodeler/goat/mountain* spoke* _of the quick response_* and he* explained* _the many lives saved_.
The yodeler/goat/mountain* answered* _all the reporter's questions_* as truthfully* as possible.
The yodeler/goat/mountain* was also polite* _to the reporter_* and did well* in the interview.

3

There once was* a magician/rabbit/hat that* owned* _a new fancy restaurant downtown_.
On opening night,* the magician/rabbit/hat* invited* _his friends, and he_* advertised* _his dishes_.
The magician/rabbit/hat* was very thrilled,* _and the night_* went smoothly.
The magician/rabbit/hat* declared,* _"Maybe this was a success,"_* as the last customer left.
This was the first time* that the magician/rabbit/hat* ventured* _into the restaurant business_.

4

On an ordinary* Monday,* a beekeeper/bee/honeycomb* strolled* _into the bank_.
The beekeeper/bee/honeycomb* requested* _a withdrawal slip_* so she could* take out $100.
The bank teller asked* for the account information.
The beekeeper/bee/honeycomb* rummaged* _in her wallet_* and she* whispered* _her information_.
The beekeeper/bee/honeycomb* was buying* _a birthday gift_* for her daughter.

5

After a bout* of sickness,* a chef/lobster/butter stick* scheduled* _a doctor's appointment_.
Luckily, an appointment was available* the very next day.
The chef/lobster/butter stick* had sneezed* _too many times_* the past week.
The chef/lobster/butter stick *was prescribed* _medication_* and he* listened* _to advice_* to get rest.
For the next week,* though, the chef/lobster/butter stick* blew* _his runny nose_* continually.

6

Across the world,* there was a diver/shark/cage that* announced* _baseball games_* in his city.
At the start of a big match-up,* the diver/shark/cage* introduced* _himself and then_* sung* _their anthem_.
After the anthem,* the diver/shark/cage* discussed* _each player_* and his statistics.
It was the fiftieth game* of the season.
The weather became ugly,* and the* diver/shark/cage* broadcasted* _a delay of game_.

7

This summer,* a trainer/dolphin/whistle* volunteered* _as a camp counselor_* in the Rockies.
The trainer/dolphin/whistle* remembered* _her favorite activities_* at the same camp* years ago.
These memories included* swinging* _on the tire_* and also* canoeing* _on the lake_.

The campers always had* a lot of fun with their counselors* at this camp.
The trainer/dolphin/whistle* prayed* _that no camper_* would get injured during the week.

8

One day in a studio,* a policeman/German shepherd/gun* was the host* _of a daytime game show_.
They were currently filming* a holiday episode.
As the contestant played*, the policeman/German shepherd/gun* gestured* _for a commercial_* break.
The policeman/German shepherd/gun* calculated* _the time and he_* commanded* _the music_* to stop.
As fake snow* fell on the stage,* the policeman/German shepherd/gun* yelled,* _"And we're back!"_

9

In a far country,* an/a Eskimo/dog/sled* taught* _geology_* at a school on a hilltop.
For fun,* the Eskimo/dog/sled* built* _a volcano and_* poured* _baking soda_* and vinegar into it.
The Eskimo/dog/sled also* ordered* _geodes_* and then he* demonstrated* _how to break them_.
The students enjoyed these lessons* very much.
And the young children* were inspired* to pursue science* as a career.

10

Last Thursday,* a farmer/cow/milk bottle* exercised* _at the gym_* down the street.
The farmer/cow/milk bottle* deadlifted* _100 pounds_* and then* jogged* _2 miles_* on the treadmill.
The leg press was another* planned workout* for the day.
Afterwards, the farmer/cow/milk bottle* wiped* _down_* all the machines.
The farmer/cow/milk bottle* crawled* _out of the gym_* exhausted.

11

Some time ago,* a zookeeper/lion/tree* bought* _tickets to an amusement park_.
The trip* had been scheduled for months.
On every spinning ride, the zookeeper/lion/tree* giggled* _uncontrollably and_* announced* _he was dizzy_.
The zookeeper/lion/tree* reminisced* _about_* how different the park looked* years ago.
The zookeeper/lion/tree* was impressed* _with all the updated_* rides and facilities.

12

At one point, a jockey/horse/saddle* went shopping* _at a large grocery store_.
The jockey/horse/saddle* was disappointed* _to find the store_* overly crowded.
The jockey/horse/saddle* muttered* _some_* not-so-nice words* under his breath.
After a couple seconds of contemplation, the jockey/horse/saddle* decided* _he would try_* another store.
The jockey/horse/saddle* drove* _out of the parking lot_* in a hurry.

13

On a highly-anticipated* movie release night,* a hunter/deer/gun* drove* _to the theater_.
The hunter/deer/gun* asked* _someone_* if there were* tickets* for the movie.
It was sold out,* so the hunter/deer/gun* chose* _another one and_* handed* _his ticket to scan_.
But inside the theater,* an alternative* presented itself.
The hunter/deer/gun* slithered* _into_* the wrong theater room* to watch* the newest movie.

14

In the olden days,* a farmer/chicken/feed barrel* was employed* _as the pharmacist_* on Saturdays.
Most of the time,* the farmer/chicken/feed barrel* counted* _medication and_* talked* _with his neighbors_.
Even though the hours were long,* it was a good job.
The farmer/chicken/feed barrel* was notorious for* gossiping* _plenty and for_* bragging* _about his farm_.
But everyone in town was happy* to have such a reliable pharmacist.

15

On Friday,* a cowboy/horse/saddle* consulted* _with his therapist_.
The therapist* was a specialist* in treating mood swings.
After the cowboy/horse/saddle* questioned* _his therapist_* for ways to control* his mood.
The cowboy/horse/saddle* laughed* _at the thought_* of exercise and he* cringed* _for the no drinking suggestion_.

The cowboy/horse/saddle begrudgingly* reflected,* *"there goes my** Sunday football routine."

16

Once a veterinarian/dog/stethoscope* was gossiping* *about a telemarketer** who frequently called.
When the phone rang again,* the veterinarian/dog/stethoscope* brainstormed* *the joke** she would play.
The veterinarian/dog/stethoscope* pretended* *to be a parrot** and she* repeated* *every word** of the caller.
Then the veterinarian/dog/stethoscope* advised* *the telemarketer** to leave a message* after the beep.
Confused,* the telemarketer hung up.

17

There once was* a man/cat/tree that* assembled* *tents for the** traveling circus.
It was a tough job* but the tent workers* could watch the circus for free.
The man/cat/tree* whistled* *noisily and** clapped* *loudly for the** elephants' performance.
When the act ended,* the man/cat/tree* cheered* *for an encore*.
The man/cat/tree* envisioned* *an even bigger** circus next year.

18

For the first time,* a dog show handler/dog/trophy* borrowed* *a book** from the library.
The book was quite long* but contained important information.
The dog show handler/dog/trophy* read* *the entire book** in two hours.
The dog show handler/dog/trophy* wrote* *a lot of notes*,* and she* googled* *for more books** to read.
The next day, the dog show handler/dog/trophy* ordered* *an Uber** to go back to the library.

19

The marine biologist/fish/boat* jogged* *into a department store** to start work.
The next shift* started* in five minutes.
Once it started, the marine biologist/fish/boat* cataloged* *new inventory** that was in boxes.
The marine biologist/fish/boat* ironed* *the clothes** and then* requested* *another project*.
The manager shook his head,* for the marine biologist/fish/boat* forgot* *to unpack** a few boxes.

20

After an interesting* morning,* a blind man/service dog/walking stick* telephoned* *his curator friend*.
The curator* had been a close friend* since high school.
The blind man/service dog/walking stick* purchased* *artwork and** inquired* *if his friend** would examine it.
After an hour, the blind man/service dog/walking stick* texted* *the curator** for his* whereabouts.
Then the blind man/service dog/walking stick* mopped* *the kitchen floor** while he* waited* for his friend.

1f

One fall day,* a barista* was hiking* with his friend.
The friend* was whistling loudly in the woods.
The barista* rolled his eyes* and then almost* tripped over a tree root.
Then,* he laughed nervously* so that his friend* glanced* at him.
The barista* muttered, "The tree almost got me."

2f

One day a radio host* was picnicking* at Yellowstone.
The hot springs* were very beautiful* and it wasn't* too cold.
The radio host* unfolded* her blanket* and arranged her food containers.
The radio host* photographed herself* in front of the hot spring.
Then, she casually* prepared* a plate and* ate* her meal in peace.

3f

One evening* a hotel manager attended a musical.
The hotel manager* gasped* when she heard the opener and she grinned* at the singer.

She knew* the song* and liked it very much.
But the hotel manager* was apathetic to the rest of the performance.
The musical* left the hotel manager* confused as she returned* to the hotel.

4f

On Wednesday,* an astronaut* was wandering around a construction site.
The astronaut* was scanning the area* for a ring he lost* the day before.
He was startled when a worker* yelled,* "Sir, this is a dangerous area."
The astronaut* shrugged his shoulders* and replied,* "I need to find* my ring."
The construction worker* quickly hid his hand* behind his back.

5f

While on his lunch break,* a crossing guard* strolled* into a shoe store.
He was really hoping they had* some loafers* for an upcoming formal event.
He flagged down the attendant* to ask about* promotional discounts.
She explained* that discounts* only applied* to teen shoes.
The crossing guard then* messaged his partner* to please buy loafers online

6f

After a long Monday,* a graffiti artist* was relaxing* in the local firehall.
She read a sign* stating that bingo* was happening that night.
The graffiti artist* judged her odds* and concluded she would play.
She had not played* since her* high school days.
The painter was eager to see the winning prize.

7f

One afternoon* a black belt* was piloting* his hot air balloon.
The black-belt* steered* his hot air balloon* toward a barley field.
Gazing over the edge,* he tallied the number of sheep* in his landing path.
"Drat,"* he sputtered,* "why do I always forget about the sheep?"
The black belt* navigated* the hot air balloon* away from the herd.

8f

Once,* a mechanical engineer* was singing* karaoke in a bar.
She was amazed* at all the song options* she could choose from on the machine.
The mechanical engineer* danced energetically* while she sang.
She motioned* to her one friend* to begin* the grand finale.
The mechanical engineer and her friend* hurdled the bar stools.

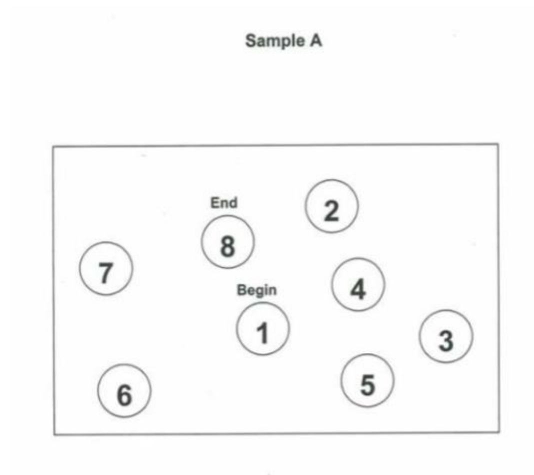
9f

Many days ago* a roller coaster enthusiast* ordered* a steak at a steakhouse.
She was dining alone* and had just* journaled about the* new wooden coaster.
Upon reflection,* the enthusiast* determined* she was no longer hungry.
When the waitress* returned,* the roller coaster enthusiast pointed to the menu.
"Can I change* my order* and just get a milkshake instead?"

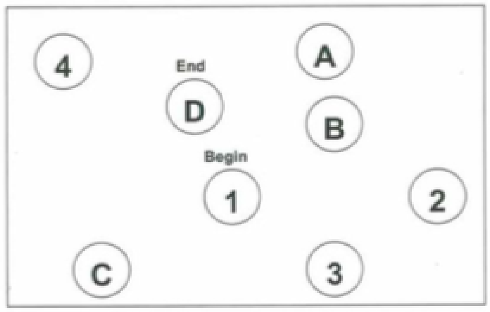
10f

In the city,* a CEO* visited* his favorite bakery.
The CEO waved to the baker* and commented,* "We're out of wheat bread again at home."
The baker* shifted* his weight uncomfortably* and said, "So are we."
The CEO* cursed* under his breath.
He gestured toward the shelf,* responding,* "I guess I'll take sourdough."

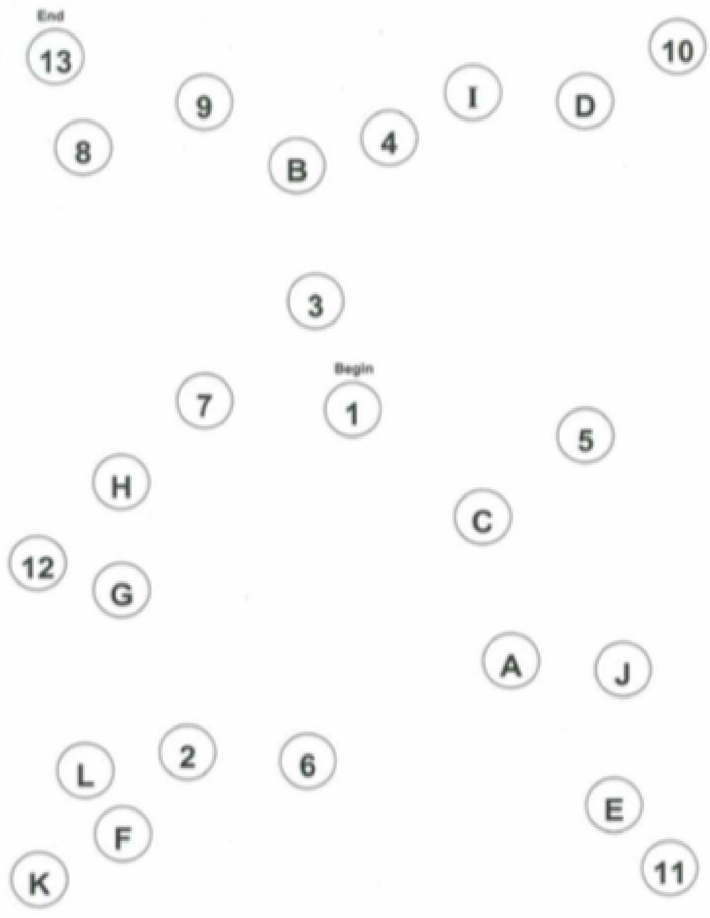
Appendix H Trails A + B Task



Sample B



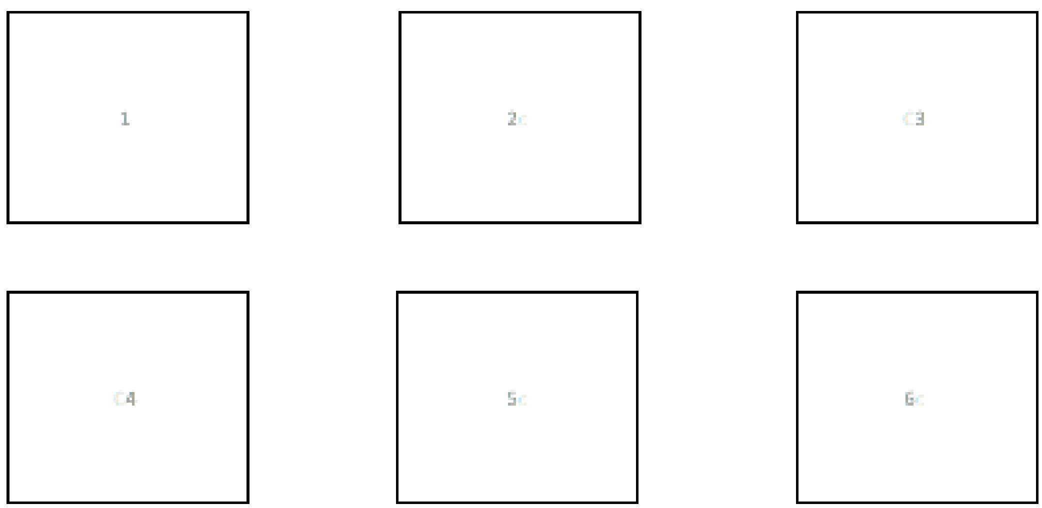
4
End
D
A
B
Begin
1
2
C
3



End
13
9
8
B
4
I
D
10
3
Begin
7
1
5
H
C
12
G
A
J
L
2
6
E
F
K
11

Appendix I Response Portion of Keep Track Task

Participants were instructed to use this paper to respond during the Keep Track task by either saying out loud the number location of the color or pointing with their finger the location of the square.



Appendix J Summary of Mixed-Effects Models for Experiment 3

Within-Narrative Adaptation

No Individual Difference Measure

Critical region

Model 1: Reading time ~ 1 + Violation Condition*Verb Position +
(1 + Violation Condition*Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.623	0.102	12.499	65.066	< 0.001
Stronger Violation (SV.) Condition	0.146	0.098	11.633	1.487	NS
Weaker Violation (WV.) Condition	-0.189	0.078	11.954	-2.421	0.032
Verb (V.) Position	-0.047	0.010	12.211	-4.779	< 0.001
SV. Condition*V. Position	-0.026	0.026	11.397	-1.018	NS
WV. Condition*V. Position	0.041	0.018	24.387	2.226	0.036

Post-critical region

Model 2: Reading time ~ 1 + Violation Condition*Verb Position +
(1 + Violation Condition*Verb Position | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	0.868	0.126	12.633	54.585	< 0.001
SV. Condition	0.130	0.125	10.543	1.038	NS
WV. Condition	-0.017	0.098	14.426	-0.176	0.863
V. Position	-0.042	0.012	11.918	-3.547	0.004
SV. Condition*V. Position	-0.008	0.031	12.201	-0.264	NS
WV. Condition*V. Position	-0.001	0.023	29.208	-0.050	NS

Shifting

Critical region

Model 3: Reading time ~ 1 + Violation Condition*Verb Position*Shifting Composite +
(1 + Violation Condition | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.624	0.076	12.590	87.479	< 0.001
SV. Condition	0.145	0.066	58.160	2.194	0.032
WV. Condition	-0.190	0.066	53.920	-2.866	0.006
V. Position	-0.047	0.006	1751.000	-7.916	< 0.001
Shifting (S.) Composite	0.220	0.099	11.250	2.218	0.048
SV. Condition*V. Position	-0.026	0.017	1751.000	-1.542	NS
WV. Condition*V. Position	0.041	0.017	1751.000	2.419	0.016
SV. Condition*S. Composite	-0.043	0.089	57.520	-0.487	NS
WV. Condition*S. Composite	0.137	0.089	53.210	1.541	NS
V. Position*S. Composite	-0.009	0.008	1751.000	-1.125	NS
SV. Condition*V. Position*S. Composite					
WV. Condition*V. Position*S. Composite					

Post-critical region

Model 4: Reading time ~ 1 + Violation Condition*Verb Position*Shifting Composite + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	t	p
Intercept	6.867	0.094	14.610	73.263	< 0.001
SV. Condition	0.133	0.070	1759.000	1.900	0.058
WV. Condition	-0.015	0.070	1759.000	-0.209	NS
V. Position	-0.042	0.007	1759.000	-5.620	< 0.001
Shifting (S.) Composite	0.288	0.118	11.380	2.437	0.032
SV. Condition*V. Position	-0.009	0.021	1759.000	-0.435	NS
WV. Condition*V. Position	-0.002	0.021	1759.000	-0.097	NS
SV. Condition*S. Composite	-0.201	0.094	1759.000	-2.151	0.032
WV. Condition*S. Composite	0.175	0.094	1759.000	1.866	0.062
V. Position*S. Composite	-0.018	0.010	1759.000	-1.851	0.064
SV. Condition*V. Position*S. Composite	0.035	0.028	1759.000	1.234	NS
WV. Condition*V. Position*S. Composite	-0.014	0.028	1759.000	0.483	NS

Inhibiting

Critical region

Model 5: Reading time ~ 1 + Violation Condition*Verb Position*Inhibiting Composite + (1 + Violation Condition | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	t	p
Intercept	6.623	0.086	11.850	75.618	< 0.001
SV. Condition	0.146	0.007	48.680	2.082	0.043
WV. Condition	-0.189	0.071	36.180	-2.646	0.012
V. Position	-0.047	0.006	1736.000	-7.924	< 0.001
Inhibiting (I.) Composite	-0.057	0.112	10.950	-0.498	NS
SV. Condition*V. Position	-0.003	0.017	1734.000	-1.556	NS
WV. Condition*V. Position	0.041	0.017	1734.000	2.416	0.016
SV. Condition*I. Composite	-0.103	0.090	47.990	-1.157	NS
WV. Condition*I. Composite	0.073	0.094	37.110	0.779	NS
V. Position*I. Composite	0.005	0.008	1738.000	0.609	NS
SV. Condition*V. Position*I. Composite	0.017	0.022	1735.000	0.795	NS
WV. Condition*V. Position*I. Composite	-0.005	0.022	1735.000	0.207	NS

Post-critical region

Model 6: Reading time ~ 1 + Violation Condition*Verb Position*Inhibiting Composite +
(1 + Violation Condition | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.867	0.107	13.310	64.138	< 0.001
SV. Condition	0.133	0.083	56.580	1.611	NS
WV. Condition	-0.014	0.084	60.770	-0.168	NS
V. Position	-0.042	0.007	1736.000	-5.660	< 0.001
I. Composite	-0.071	0.133	1736.000	-0.537	NS
SV. Condition*V. Position	-0.009	0.021	1733.000	-0.439	NS
WV. Condition*V. Position	-0.002	0.021	1733.000	-0.098	NS
SV. Condition*I. Composite	-0.041	0.108	56.600	-0.381	NS
WV. Condition*I. Composite	-0.032	0.108	59.090	0.295	NS
V. Position*I. Composite	-0.003	0.010	1737.000	-0.347	NS
SV. Condition*V. Position*I. Composite	0.007	0.027	1734.000	0.273	NS
WV. Condition*V. Position*I. Composite	0.002	0.027	1734.000	0.057	NS

WM Updating

Critical region

Model 7: Reading time ~ 1 + Violation Condition*Verb Position*WM Updating Composite +
(1 + Violation Condition | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.622	0.087	11.840	76.493	< 0.001
SV. Condition	0.148	0.070	48.750	2.096	0.041
WV. Condition	-0.188	0.072	36.630	-2.619	0.013
V. Position	-0.047	0.006	1736.000	-7.931	< 0.001
WM Updating (WM.) Composite	0.134	0.113	10.920	1.178	NS
SV. Condition*V. Position	-0.027	0.017	1733.000	-1.602	NS
WV. Condition*V. Position	0.040	0.017	1733.000	2.406	0.016
SV. Condition*WM. Composite	-0.291	0.092	47.600	-3.154	0.003
WV. Condition*WM. Composite	0.094	0.097	37.700	0.977	NS
V. Position*WM. Composite	-0.022	0.008	1734.000	-2.732	0.006
SV. Condition*V. Position*WM. Composite	0.079	0.022	1732.000	3.526	< 0.001
WV. Condition*V. Position*WM. Composite	-0.010	0.022	1732.000	-0.440	NS

Post-critical region

Model 8: Reading time ~ 1 + Violation Condition*Verb Position*WM Updating Composite + (1 + Violation Condition | Subject) + (1 + Violation Condition | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.866	0.109	13.300	63.108	< 0.001
SV. Condition	0.134	0.082	58.360	1.639	NS
WV. Condition	-0.014	0.084	58.850	-0.161	NS
V. Position	-0.042	0.007	1736.000	-5.650	< 0.001
WM. Composite	0.051	0.138	10.990	0.369	NS
SV. Condition*V. Position	-0.010	0.021	1733.000	-0.460	NS
WV. Condition*V. Position	-0.002	0.021	1733.000	-0.110	NS
SV. Condition*WM. Composite	-0.249	0.111	58.180	-2.245	0.029
WV. Condition*WM. Composite	0.143	0.111	58.600	1.289	NS
V. Position*WM. Composite	-0.019	0.010	1734.000	-1.900	0.058
SV. Condition*V. Position*WM. Composite	0.067	0.028	1732.000	2.403	0.016
WV. Condition*V. Position*WM. Composite	-0.035	0.028	1732.000	-1.260	NS

Common CC

Critical region

Model 9: Reading time ~ 1 + Violation Condition*Verb Position*Common CC Composite + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.519	0.087	13.370	75.092	< 0.001
SV. Condition	0.200	0.065	1759.000	3.059	0.002
WV. Condition	-0.215	0.065	1760.000	-3.291	0.001
V. Position	-0.037	0.007	1760.000	-5.386	< 0.001
Common CC (CC.) Composite	0.538	0.210	15.470	2.560	0.021
SV. Condition*V. Position	-0.043	0.020	1759.000	-2.195	0.028
WV. Condition*V. Position	0.042	0.020	1761.000	2.151	0.032
SV. Condition*CC. Composite	-0.284	0.170	1760.000	-1.669	0.095
WV. Condition*CC. Composite	0.123	0.170	1760.000	0.723	NS
V. Position*CC. Composite	-0.050	0.018	1760.000	-2.757	0.006
SV. Condition*V. Position*CC. Composite	0.088	0.051	1759.000	1.724	0.085
WV. Condition*V. Position*CC. Composite	-0.005	0.051	1761.000	-0.099	NS

Post-critical region

Model 10: Reading time ~ 1 + Violation Condition*Verb Position* Common CC Composite + (1 | Subject) + (1 | Item)

Predictor	Coefficient	SE	df	*t*	p
Intercept	6.761	0.117	13.070	57.677	< 0.001
SV. Condition	0.190	0.081	1759.000	2.384	0.019
WV. Condition	-0.047	0.081	1759.000	-0.583	NS
V. Position	-0.028	0.009	1759.000	-3.314	< 0.001
CC. Composite	0.546	0.024	1115.000	1.868	0.088
SV. Condition*V. Position	-0.028	0.024	1759.000	-1.137	NS
WV. Condition*V. Position	0.005	0.024	1759.000	0.210	NS
SV. Condition*CC. Composite	-0.299	0.210	1759.000	-1.423	NS
WV. Condition*CC. Composite	0.162	0.210	1759.000	0.769	NS
V. Position*CC. Composite	-0.691	0.022	1759.000	-3.079	0.002
SV. Condition*V. Position*CC. Composite	0.097	0.063	1759.000	1.533	NS
WV. Condition*V. Position*CC. Composite	-0.035	0.063	1759.000	-0.553	NS

Bibliography

Allen, C. M., Martin, R. C., & Martin, N. (2012). Relations between short-term memory deficits, semantic processing, and executive function. *Aphasiology, 26*(3-4), 428-461.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition, 114*(3), 455-461.

Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *Journal of memory and language, 63*(4), 489-505.

Bohan, J., Leuthold, H., Hijikata, Y., & Sanford, A. J. (2012). The processing of good-fit semantic anomalies: An ERP investigation. *Neuropsychologia*, *50*(14), 3174-3184.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624.

Boudewyn, M. A., Blalock, A. R., Long, D. L., & Swaab, T. Y. (2019). Adaptation to animacy violations during listening comprehension. *Cognitive, Affective, & Behavioral Neuroscience,* 1-12.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience & Biobehavioral Reviews*, *26*(7), 809-817.

van den Broek, P., Risden, K., & Husbeye-Hartmann, E. (1995). The role of readers' standards of coherence in the generation of inferences during reading. In R. F. Lorch, Jr., & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology, 23*(2), 201-216. https://doi.org/10.1016/j.acn.2007.08.010

Chang, F., Dell, G., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234-272. https://doi.org/10.1037/0033-295X.133.2.234.

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass, 6*(5), 259-278.

Chomsky, D. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Colvin, M., & Warren, T. (2020). Language comprehension models at different levels of explanation. *Language and Linguistics Compass*, *14*(1), e12362.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1-2), 26–49. https://doi.org/10.1080/0163853X.2013.855107

Cook, A. E., Walsh, E. K., Bills, M. A., Kircher, J. C., & O'Brien, E. J. (2018). Validation of semantic illusions independent of anomaly detection: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *71*(1), 113–121. https://doi.org/10.1080/17470218.2016.1264432

Creer, S. D., Cook, A. E., & O'Brien, E. J. (2018). Competing activation during fantasy text comprehension. *Scientific Studies of Reading*, *22*, 308–320.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168. https://doi.org/10.1146/annurev-psych-113011-143750.
Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179-211.

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*(APR) https://doi.org/10.3389/fpsyg.2014.00335

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current directions in psychological science*, *27*(6), 443-448.

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition*, *106*, 1038-1046. https://doi.org/10.1016/j.cognition.2007.04.006

Filik, R., & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology*, *45*, 554–558. https://doi.org/10.1111/j.14698986.2008.00656.x

Fine, A. B., & Florian Jaeger, T. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*(3), 578-591.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PloS one*, *8*(10), e77661.

Foy, J., & Gerrig, R. (2014). Flying to Neverland: How readers tacitly judge norms during comprehension. *Memory & Cognition*, *42*, 1250-1259. https://doi.org/10.3758/s1342101404368

Fraundorf, S. H., & Jaeger, T. F. (2016). Readers generalize adaptation to newly-encountered dialectal structures to other unfamiliar structures. *Journal of Memory and Language*, *91*, 28-58.

Frazier, L. (1987). Theories of sentence processing. In J. L. Garfield (Ed.) *Modularity in knowledge representation and natural-language understanding* (pp. 291-307). Cambridge, MA: MIT Press.

Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39*(2-3), 225–242. https://doi.org/10.1080/0163853X.2005.9651681

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, *110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*(4), 263–268.

Hagoort, P. Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438-441. https://doi.org/10.1126/science.1095455.

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological science*, *27*(4), 572-582.

Jaeger, T. F., & Snider, N. (2008, July). Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proceedings of the cognitive science society conference* (pp. 1061-1066). Washington, DC: Cognitive Science Society.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57-83.

Kamas, E. N., & Reder, L. M. (1995). The role of familiarity in cognitive processing.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*(2), 170-210. https://doi.org/10.2307/411200

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182.

Kuperberg, G. R. (2016). Separate streams or probabilistic inference? What the N400 can tell us about the comprehension of events. *Language, Cognition, and Neuroscience*, *31*(5), 602–616. https://doi.org/10.1080/23273798.2015.1130233

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, cognition and neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Loucks, J., Verrett, K., & Reise, B. (2020). Animates engender robust memory representations in adults and young children. *Cognition*, *201*. https://doi.org/10.1016/j.cognition.2020.104284

Marr, D. (1982). *Vision: A Computational Approach*. San Francisco: Freeman & Co.

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(4), 913-934. https://doi.org/10.1037/a0022964

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543-562.

McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: Some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22* (4), 919.

McKoon, G., & Ratcliff, R. (1998). Memory-based language processing: Psycholinguistic research in the 1990s. *Annual Review of Psychology*, *49*(1), 25–42. https://doi.org/10.1146/annurev.psych.49.1.25

McRae, K., & Matsuki, K. (2009). People use their knowledge of common events to understand language, and do so as quickly as possible. *Language and Linguistics Compass*, *3*(6), 1417-1429. https://doi.org/10.1111/j.1749-818X.2009.00174x

Milburn, E., Warren, T., & Dickey, M. W. (2016). World knowledge affects prediction as quickly as selectional restrictions: Evidence from the visual world paradigm. *Language, Cognition, and Neuroscience*, *31*(4), 536-548. https://doi.org/10.1080/23273798.2015.1117117

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current directions in psychological science*, *21*(1), 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2-3), 131–157. https://doi.org/10.1080/01638539809545042

Nieuwland, M., & van Berkum, J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.

Novick, JM, Hussey, E, Teubner-Rhodes, S, Harbison, JI, & Bunting, MF (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition, and Neuroscience*, 29(2), 186-217.

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive neuropsychology*, *26*(6), 527-567. https://doi.org/10.1080/02643290903519367

Novick, JM, Trueswell, JC, & Thompson-Schill, SL (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, 4(10), 906-924.

O'Brien, E. J., & Cook, A. E. (2016a). Separating the activation, integration, and validation components of reading. In *Psychology of Learning and Motivation* (Vol. 65) (pp. 249–276). Academic Press.

O'Brien, E. J., & Cook, A. E. (2016b). Coherence threshold and the continuity of processing: The RI-Val model of comprehension. *Discourse Processes*, *53*(5-6), 326–338. https://doi.org/10.1080/0163853X.2015.1123341

Pylkkänen, L., Oliveri, B., & Smart, A. J. (2009). Semantics vs. world knowledge in prefrontal cortex. *Language and Cognitive Processes*, *24*(9), 1313-1334. https://doi.org/10.1080/01690960903120176

R Core Team (2014). *R: A language and environment for statistic computing* [Computer software]. 536 Vienna, Austria: R Foundation for Statistic Computing. https://www.Rproject.org/.

Rayner, K., Warren, T., Juhasz, B., & Liversedge, S. (2004). The effects of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1290-1301. https://doi.org/10.1037/0278-7393.30.6.1290

Ryskin, R., Levy, R. P., & Fedorenko, E. (2020). Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. *Neuropsychologia*, *136*, 107258. https://doi.org/10.1016/j.neuropsychologia.2019.107258

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime: User's guide.* Psychology Software Incorporated.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition, 89*(3), 179-205.

Sha, L., Haxby, J. V., Abdi, H., Guntupalli, J. S., Oosterhof, N. N., Halchenko, Y. O., & Connolly, A. C. (2015). The animacy continuum in the human ventral vision pathway. *Journal of cognitive neuroscience*, *27*(4), 665-678.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, *138*, 107307. https://doi.org/10.1016/j.neuropsychologia.2019.107307

Simic, T. (2019). Predictors of post-stroke aphasia recovery (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.

Simic, T., Bitan, T., Turner, G., Chambers, C., Goldberg, D., Leonard, C. & Rochon, E. (2019). The role of executive control in post-stroke aphasia treatment. *Neuropsychological Rehabilitation*, 1-40. https://doi.org/10.1080/09602011.2019.1611607

Stuss, D.T. & Alexander, M.P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of The Royal Society Biological Sciences*, *362*, 901 – 915.

Sutton, R. S., and Barto, A. G. (1981). Toward a modern theory of adaptive networks – expectation and prediction. *Psychol. Rev.* 88, 135–170. https://doi.org/10.1037/0033-295x.88.2.135

Takiyama, K., Hirashima, M., & Nozaki, D. (2015). Prospective errors determine motor learning. *Nature communications*, *6*(1), 1-12.

Tanaka, S., O'Doherty, J. P., & Sakagami, M. (2019). The cost of obtaining rewards enhances the reward prediction error signal of midbrain dopamine neurons. *Nature communications*, *10*(1), 1-13.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Towards a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Vuong, L. C., & Martin, R. C. (2011). LIFG-based attentional control and the resolution of lexical ambiguities in sentence context. *Brain and Language*, *116*(1), 22-32.

Vuong, L. C., & Martin, R. C. (2014). Domain-specific executive control and the revision of misinterpretations in sentence comprehension. *Language, Cognition and Neuroscience*, *29*(3), 312-325.

Walsh, E. K., Cook, A. E., & O'Brien, E. J. (2018). Processing real-world violations embedded within a fantasy-world narrative. *Quarterly Journal of Experimental Psychology*, *71*(11), 2282–2294.

Warren, T. (2011). *The influence of implausibility and anomaly on eye movements during reading*. The Oxford handbook of eye movements: Oxford University Press.

Warren, T., & Dickey, M. W. (2020). *The use of linguistic and world knowledge in language processing.* Revised manuscript submitted for publication.

Warren, T., Dickey, M. W., & Liburd, T. L. (2017). A rational inference approach to group and individual-level sentence comprehension performance in aphasia. *Cortex*, *92*, 19–31. https://doi.org/10.1016/j.cortex.2017.02.015

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about plausible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(4), 1001-1010. https://doi.org/10.1037/0278-7393.34.4.1001

Warren, T., Milburn, E., Patson, N. D., & Dickey, M. W. (2015). Comprehending the impossible: What role do selectional restriction violations play? *Language, Cognition, and Neuroscience, 30*(8), 1-8. https://doi.org/10.1080/23273798.2015.1047458

Williams, C. R., Cook, A. E., & O'Brien, E. J. (2018). Validating semantic illusions: Competition between context and general world knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1414.

Willits, J. A., Amato, M. S., & MacDonald, M. C. (2015). Language knowledge and event knowledge in language use. *Cognitive Psychology, 78,* 1-27. https://doi.org/10.1016/j.cogpsych.2015.02.002