

2013

# The Role of World Knowledge and Episodic Memory in Scripted Narratives

Micah L. Mumper

Macalester College, micah.mumper@gmail.com

Follow this and additional works at: [http://digitalcommons.macalester.edu/psychology\\_honors](http://digitalcommons.macalester.edu/psychology_honors)



Part of the [Cognitive Psychology Commons](#)

---

## Recommended Citation

Mumper, Micah L., "The Role of World Knowledge and Episodic Memory in Scripted Narratives" (2013). *Psychology Honors Projects*. Paper 32.

[http://digitalcommons.macalester.edu/psychology\\_honors/32](http://digitalcommons.macalester.edu/psychology_honors/32)

This Honors Project is brought to you for free and open access by the Psychology Department at DigitalCommons@Macalester College. It has been accepted for inclusion in Psychology Honors Projects by an authorized administrator of DigitalCommons@Macalester College. For more information, please contact [scholarpub@macalester.edu](mailto:scholarpub@macalester.edu).

The Role of World Knowledge and Episodic Memory in Scripted Narratives

Micah Mumper

Macalester College

Authors Note:

Micah Mumper, Department of Psychology, Macalester College.

I am indebted to R. Brooke Lea, Darcy Burgund, and Christina Esposito for their helpful comments on this article. I would also like to thank Anne Cook, Jerry Myers, and Kristin Weingartner for making their materials available. Finally, thank you to my family and all my friends who supported me throughout this process.

Correspondences concerning this article should be sent via email to micah.mumper@gmail.com.

### Abstract

Readers recruit information from both general world knowledge and episodic memory during reading comprehension. The present experiment used eye tracking to investigate the time-course of how these two sources of memory interact. Participants read passages describing scenarios in which an actor performs a role that was either scriptually appropriate or inappropriate. Half the passages containing the inappropriate role-filler were preceded by an episodic justification for this scriptural violation. Using the same paradigm, Cook and Myers (2004) found context had an early influence on the integration of the role-filler, but world knowledge showed a later effect in the post-target region. The present experiment expands upon these results by adding a backgrounding section between the episodic justification and the role-filler that reduces the saliency of the episodic information. Evidence of integration difficulties were observed in both early and late processing measures. These results appear to favor a two-stage model of processing: the first stage links incoming information with the contents of active memory; the second stage evaluates the link between the old and new information.

### The Role of World Knowledge and Context in Scripted Narratives

Most theories of text comprehension assume that the basic goal of reading is to establish a coherent representation of the text. In doing so, readers construct an elaborate mental representation of what the text is about, known as the situation model (Gueraud, 2005; Kintsch, 1988; Zwaan & Radvansky, 1998). For successful comprehension, the situation model must be updated to include incoming information from the text. However, due to the limited amount of information that can be processed at once, only a select portion of the situation model is active in working memory for updating (e.g., Albrecht & O'Brien, 1993). Typically, to facilitate the integration of incoming information, readers maintain in working memory the immediately preceding context (i.e., generally the previous one to three sentences) and related information reactivated from long-term memory (e.g., Albrecht & O'Brien, 1993; Cook, Halleran, & O'Brien, 1998; Hakala & O'Brien, 1995). This reactivated information from long-term memory includes contextual information from backgrounded portions of a text (i.e., episodic memory) that are either necessary to maintain coherence or simply related to the information currently active in memory, as well as related information from general world knowledge (i.e., semantic memory) to fill in details that are not explicitly stated in the text. Theories differ, however, in how reactivated episodic and semantic information are managed by readers as they attempt to build this mental representation. The focus of the present investigation is to understand the time-course of how these sources of information become integrated into the situation model, and the influences they have on reading comprehension.

Some discourse processing theories assume that general world knowledge has an earlier effect on processing than episodic information. For example, Kintsch's (1988) construction-integration model incorporates both immediate effects due to the semantic properties of the text and later effects due to discourse context. In the construction stage, propositions and concepts from the reader's semantic memory are activated as a function of their relatedness to what is currently in the readers focus (e.g., working memory); prior discourse context does not exert an effect until the later integration stage.

Similarly, Garrod and Terras (2000) have provided evidence that general world knowledge has an immediate effect on language processing. They tracked readers' eye-movements for passages containing a common verb (e.g., writing) in a specific scenario (e.g., writing a letter; writing on the blackboard). After introducing the scene, the following sentence contained a reference to the role-filler that was either: explicitly stated or implied; dominant (e.g., pen) or non-dominant (e.g., chalk) with respect to the verb (e.g., writing); and either appropriate (e.g., writing a letter with a pen) or inappropriate (e.g., writing on the blackboard with a pen) with respect to the described scenario. The next sentence contained the target reference to the role-filler (e.g., The pen/chalk dropped to the floor). It was found that if the role-filler was non-dominant with respect to the verb (e.g., chalk), then initial processing times on the target role-filler were slower when the role-filler was not explicitly stated, regardless of the appropriateness of the context in which the role-filler was introduced. This was taken as evidence for an early effect of general world knowledge. The appropriateness of the role-filler did not influence processing until

further downstream, suggesting that episodic factors had a late effect of processing. Garrod and Terras explained their results using a two-stage bonding and resolution model. In the bonding stage, an initial link is established between the role-filler and general world knowledge reactivated via a bottom-up memory retrieval process. This link is then evaluated and resolved with respect to the episodic context during the resolution stage. Thus, in the bonding stage, readers establish the link between the role-filler and the verb, and experience processing difficulties when the role-filler is non-dominant, as general world knowledge leads readers to expect the dominant role-filler. Discourse context does not affect processing until further downstream when readers try to resolve the conflict of the teacher writing on the blackboard with a pen in the resolution stage.

World knowledge also has a persistent effect on processing. For example, the subordinate bias effect (e.g., Rayner, Pacht, & Duffy, 1994) refers to the finding that when a reader encounters an ambiguous word for which one meaning is more dominant than the other meaning (e.g., “bank”), when the prior episodic context supports the dominant meaning (e.g., “money”), fixation times on the word are facilitated compared to a non-ambiguous control condition. However, when the context supports the subordinate meaning (e.g., “river”), fixation times are longer compared to control condition. Thus, even though context supports only the subordinate meaning of the word, readers still experience processing difficulties, suggesting that the cognitive system prioritizes the semantic information over context (i.e., the dominant meaning of “bank” takes priority to the textual context that favors the river-side meaning).

A second set of theories place more emphasis on the discourse context rather than world knowledge. These are the lexical access models of Foss and colleagues, which assumes that lexical associations mediate early access to the inactive portion of the discourse model (Foss, 1982; Traxler & Foss, 2000), such as the lexical reinterpretation hypothesis (Hess et al., 1995). According to the lexical reinterpretation hypothesis, when a concept is first processed all related information reactivated from world knowledge becomes available, through a bottom-up retrieval process. However, the context in which the word is encountered modifies the representation of that concept. When that concept is encountered further along in the text, the modified representation is reactivated, facilitating processing for that concept even if it originally was inconsistent with world knowledge.

Hess et al., (1995) used a cross-modal priming procedure, in which readers listened to sentences, and were asked to pronounce the final word of the last sentence out loud when it was presented on a computer screen in front of them. It was assumed that the speed of pronunciation reflected processing ease. When readers encountered the isolated sentence “The English major wrote the poem,” their ability to pronounce the final word *poem* was facilitated compared to the neutral sentence “The boy read the poem.” However, if the reader was first given background context about the English major being in a computer science course, their representation of what it means to be an English major in this context appeared to be modified to include the “student in a technical course on computing” information. When the reader later encountered the sentence “The English major

wrote the poem,” no facilitation was observed. Thus, the contextual information appeared to modify readers’ representations of the English major to include the computer science information. Accordingly, the target word, *poem*, was not readily integrated into this representation and therefore received no facilitation. Similarly, when the sentence “The computer science major wrote the poem” was read in isolation, the word *poem* was pronounced more slowly compared to the control sentence. However, when that sentence was preceded by context that placed the computer science major in an English course, the pronunciation of the final word *poem* was facilitated. Thus, even though general world knowledge about computer science majors should have disrupted processing, the preceding discourse context appeared to override its influence by updating reader’s representation of a computer science major. It is important to recognize the limitations of the probe technique used by Hess et al. (1995). This method only provides a snapshot into the time-course of processing, and thus may not provide the complete view of processing that can be captured by methods such as eye tracking. For example, if Garrod and Terras (2000) only probed for the target word *pen*, they may have concluded that world knowledge was the only factor affecting processing, and missed the effect of context that occurred downstream. Instead, they were able to find evidence that showed world knowledge had an early effect on processing, and context did not have an effect until subsequent processing of the target word.

Nevertheless, the lexical reinterpretation hypothesis was supported by an eye-tracking experiment by Duffy and Keir (2004). Their experiment found that disambiguating the gender of a role-violating stereotype (e.g., a female electrician)



slowed reading time on the first disambiguating presentation of the role-filler, but reading times were faster on a later gendered target pronoun compared to a condition in which the gender was not disambiguated. Reading times in the disambiguated condition were not statistically different from the condition when the stereotype was not violated, suggesting that the disambiguating episodic information facilitated comprehension despite the conflicting general world knowledge about stereotypes. In the post-pronoun region, stereotype mismatch produced longer reading times only when the stereotype was not disambiguated. Thus, these results suggest that when the gender was disambiguated, readers updated their mental models to include the stereotype violating information. Later references to the violating stereotype did not produce any observable processing difficulties, in line with the hypotheses of the lexical reinterpretation hypothesis.

Other experiments have also provided evidence in favor of context having an early effect on processing. Wiley and Rayner (2000) demonstrated that initial reading times for lexically ambiguous words with more than two meanings (e.g., “fly”) were facilitated compared to non-ambiguous control words when they were embedded in texts preceded by titles that biased the word towards a particular meaning (e.g., *Worries of a Baseball Manager*; *Insects that Spread Diseases*). Participants read passages that had two distinct interpretations, depending on which of two titles were given. Critically, embedded within each passage were ambiguous words (or control words) that allowed the same passages to have alternate interpretations. Depending on the context, the correct meaning of the ambiguous nouns was subordinate, balanced, or dominant compared to the other

possible meanings. They measured reading times for the target words as well as the number of regressions. Compared to control words, there was no difference in reading times or number of regressions for all but the least frequent meaning of the ambiguous words. Thus, the presence of a title was sufficient to eliminate the usual subordinate bias effect for all but the least frequent meanings of ambiguous words. In addition, Myers, Cook, Kambe, Mason, and O'Brien (2000) manipulated the semantic and episodic availability of information needed to comprehend a sentence. Participants read passages in which a critical sentence pair was embedded. The first sentence presents a category (e.g., emergency supplies), and the second sentence provides an exemplar from that category that was either typical (e.g., candles) or atypical (e.g., lantern). To connect the two sentences, participants must create a bridging inference between the category and the exemplar (e.g., the candles/lantern were part of the emergency supplies). When presented in isolation, the second sentence was read more quickly when it contained a typical exemplar. However, when the exemplar was highly elaborated several sentences earlier, the ordinarily strong semantic effects from typicality were no longer observed. Another experiment using ERP methodology found that a biasing sentence context facilitated processing on lexically ambiguous words early in the reading process (Sereno, Brewer, & O'Donnell, 2003). Given these results, it is conceivable that episodic information may influence processing as early as general world knowledge.

O'Brien and his colleagues have provided convincing evidence that episodic information provided earlier in a text is automatically reactivated when the reader encounters related information; this reactivation occurs regardless of whether the

reactivated information facilitates or hinders processing of the current information (e.g., Albrecht & O'Brien, 1993; Myers & O'Brien, 1994; O'Brien, Rizzella, Albrecht, & Halleran, 1998). Given that the same activation mechanism (e.g., resonance, see Myers & O'Brien, 1998) is assumed to operate on the episodic memory trace and general world knowledge in parallel, there ought to be conditions when reactivated world knowledge facilitates comprehension and conditions in which it hinders comprehension. For example, imagine you are reading a story about a rock band when you encounter a sentence about a contract being signed by the band's manager (Cook & Myers, 2004). Such a scenario is consistent with your general world knowledge about the role of managers in rock bands. In this case, your general world knowledge would facilitate the integration of the manager role into the situation model. However, imagine the band's guitarist signs the contract rather than the manager. Reactivated general world knowledge about the role of guitarists in a rock band script will likely conflict with the action performed in this scenario. If it does, comprehension would be hindered, and you might be expected to take a little longer to read the word *guitarist* while you try to resolve that discrepancy. Now, imagine that one sentence prior you had read that some of the band members handled publicity and finances for the band in addition to performing. In the next sentence, when you read about the guitarist signing the contract, you have access to both the world knowledge about the appropriateness of the role-filler as well as the contextual justification for why the typical roles might be violated. How might these conflicting sources of information influence the integration of the role-filler into the situation model? Would this episodic information be enough to overpower the

script-based expectation, or would general world knowledge still exert an observable effect on processing?

Investigating how readers process texts when the discourse context conflicts with general world knowledge motivated an experiment conducted by Cook and Myers (2004, Experiment 2). They gave readers short passages about well-known scripts (e.g., a rock band). Near the end of the passage, a scripted action (e.g., playing a song or signing a contract) was performed by a role-filler (e.g., guitarist) who was either scriptually appropriate or inappropriate. In half the passages, the scripted action was justified a sentence prior (e.g., some of the band members handled publicity and finances in addition to performing), or the text was neutral with regards to the scripted action. This created a 2 x 2 design in which appropriateness (appropriate or inappropriate role filler) was crossed with the justification (present or neutral). Using eye-tracking methodologies, Cook and Myers (2004) found that readers' initial reading time for the role-filler (guitarist) in the inappropriate neutral condition were longer than reading times in both of the appropriate conditions (which were not significantly different from each other), suggesting an "appropriateness" effect wherein the reactivated information from general world knowledge disrupted comprehension. This disruption, however, disappeared when the inappropriate action was previously justified, such that their initial reading times of the target word in the inappropriate justified and both appropriate conditions were similar. This suggests that episodic information contained in the justification dominated the initial processing of the text over general world knowledge about the appropriateness of the role-filler. However, evidence for

processing difficulties was observed in the post-target measures of both inappropriate conditions (compared to the appropriate conditions), suggesting that the relevant world knowledge was reactivated, but the presence of the contextual justification delayed the integration of this information. Thus, although there was no appropriateness effect for the initial reading times when the target had been justified, measures of subsequent processing found processing difficulties in integrating the inappropriate role filler. Cook and Myers (2004) reasoned that the preceding episodic context was strong enough for readers to initially accept the inappropriate role filler, but integration of the role-filler into the reader's situation model was delayed when it was checked against world knowledge.

To explain these results, Cook and Myers (2004) presented a modified version of the two-stage bonding and resolution model (Garrod & Terras, 2000). In the first stage of the updated model, the contents of working memory are linked with the first information reactivated from long-term memory via a bottom up retrieval process (e.g., resonance). However, unlike the bonding and resolution model, which assumed that general world knowledge was always reactivated before episodic information, the model of Cook and Myers posits that the initial link can be established with either semantic or episodic information. Which type of information, if either, dominates this initial stage of processing depends on factors such as the degree of featural overlap between the contents of active memory and long-term memory, as well as the types of episodic manipulations used. Thus, during this early stage, both types of information may be reactivated and could interact to influence processing. During the second stage, the initial linkages are

evaluated and verified with additional information that continues to be reactivated and integrated with the contents of working memory. The stages are assumed to be overlapping and continuous, such that even after the eyes moves on in the text, information may continue to be verified, even as new links are made. Thus, processing difficulties may not manifest until later downstream. Cook and Myers (2004) argued that although general world knowledge often becomes available more quickly than episodic information and thus has an early influence on the initial linkage stage, under certain conditions, strong episodic information can be reactivated quickly enough to influence early processing. As information (semantic or episodic) continues to be reactivated and integrated with the incoming text, those effects appear in the second, slower verification stage.

The present experiment attempts to test the two-stage model proposed by Cook and Myers (2004) against the lexical reinterpretation hypothesis (Hess et al., 1995). To accomplish this, the availability of the contextual justification used in the passages used by Cook and Myers (2004) was decreased, so that this information is no longer present in working memory, but must be reactivated along with the general world knowledge from the reader's episodic memory of the text. This was achieved by adding a section between the presentation of the episodic justification and the role-filler that backgrounds the justification material. Both theories would predict that in the inappropriate neutral condition, reading times for the target word should be significantly slower compared to an appropriate neutral control condition, due to the conflict with general world knowledge. The critical test between the theories was between the control condition and the inappropriate

justified condition. According to the two-stage theory, the longer backgrounding section should serve to reduce the saliency of the contextual justification, such that when the inappropriate role-filler is encountered, the episodic trace would not be reactivated as quickly as it was in Cook and Myers' original experiment. Thus, this information must be reactivated from long-term memory along with general world knowledge. When the reader encounters the inappropriate role-filler, the reactivated information from both sources may interact to influence processing in the initial linkage stage, resulting in a significant slow-down for the initial reading times for the target word in the inappropriate justified condition. Alternatively, according to the lexical reinterpretation hypothesis, the added backgrounding section should have no effect on reading times. This is because their model predicts that the contextual justification would become integrated into the situation model, and thus the justification should override the influence of general world knowledge regardless of the distance in the text.

The theories also make different predictions about the possible post-target effects that may be observed. According to the two-stage model, as the initial links continue to be evaluated in the second stage, the availability of the contextual justification should allow the reader to resolve the coherence break more quickly than if this justification was not present. Thus, in terms of post-target measures, the two-stage model would predict the most post-target processing difficulties for subjects in the inappropriate neutral condition, followed by the inappropriate justified condition, and finally the least amount of difficulties in the appropriate neutral condition. The lexical reinterpretation model would only predict greater

processing difficulties in the post-target region for the inappropriate neutral condition compared to the other two conditions, which should not be significantly different from each other. This would be predicted because according to the lexical reinterpretation model, the contextual justification should be sufficient to allow the integration of the inappropriate role-filler, so the reader should not experience any further processing difficulties.

## **Method**

### **Participants**

Thirty-seven native English-speaking students with normal or corrected-to-normal vision from Macalester College participated as part of a requirement for an introductory psychology course or were recruited through word of mouth. One participant was removed because their native language was not English, and another five participants were removed due to calibration errors or failure to follow instructions. This left 31 participants in total.

### **Materials and Design**

The materials used were 24 passages adapted from Cook and Myers (2004; see Appendix A for an example passage). The first sentence introduced the script. The next sentence served as a context sentence that either justified a later action or was neutral with respect to the actions and role fillers, but continued the story. For example, the context section in the example passage in Appendix A presents information that either described some of the tenants as staff living in the apartment complex (justified condition) or described the tenants as satisfied with the condition of the building (neutral condition). The next two or three sentences



served as background. These sentences were not present in the original materials, but were written exclusively for this experiment. The backgrounding materials do not refer to either the contextual information that precedes it, or the actions of the role-filler that follow. A target sentence followed the backgrounding section, containing a scripted action completed by an appropriate or inappropriate role-filler, and several more words before the end of the sentence. It is important to note that this was the first encounter with the role-filler in the text. Reading times were measured for the role-filler as well as the post-target region of 2 or 3 words after the target role-filler word. The passage concluded with a final sentence. Thus, context was crossed with appropriateness to create three between-subjects conditions: appropriate neutral (AN), inappropriate neutral (IN) and inappropriate justified (IJ). The original experiment by Cook and Myers (2004) included an appropriate justified (AJ) condition, but because no significant differences were found between the AJ and AN condition and the present experiment makes no specific predictions about a difference between these two conditions, the AJ condition was excluded for simplicity.

In addition to the 24 experimental passages, there were 24 filler passages and two practice passages. The participants were randomly assigned to condition, and the 48 passages in each condition were pseudo-randomly presented to the participants, such that no more than three experimental passages were presented in a row.

### **Procedure and Apparatus**

Eye movements were monitored by an SR Research Ltd. (Ottawa, Ontario, Canada) EyeLink 2000 eye tracking system. The eye-tracker is a video-based tracking system with a camera and infrared LEDs that illuminate the eyes. The camera samples pupil location at a rate of 1000 Hz (a temporal resolution of 1 ms). Although viewing is binocular, following Cook and Myers (2004) procedure, only data from the right eye was recorded and used for analysis. The paragraphs were presented in lowercase letters (except when capitals were appropriate) on an 18-in. AccuSync CRT color monitor with an 85 Hz refresh rate.

After signing the consent form, the participant was set into the chin rest and asked to remain in the same position for the duration of the experiment. The eye-tracker was then calibrated to the participant's eye before beginning the practice trials. To correct for minor head movements, each trial began with a drift correction, the location of which coincided with the location of the first letter of the first word in the paragraph. When the participant's gaze was directed at the target, the experimenter triggered the appearance of the passage. The entire passage was presented on the screen at once, and participants read at a self-directed pace.

Before beginning the experiment, the participants were instructed to read for comprehension in order to be able to correctly answer a comprehension question that appeared after some of the passages. To ensure that the passages were being read for meaning, approximately 25% of the 48 passages were followed by a two-alternative-forced choice comprehension question, of which there was an equal amount of "yes" and "no" correct answers. Similar to Cook and Myers (2004), all participants answered at least 85% of these questions correctly, and, following their

procedure, the fixation data for all trials were analyzed, regardless of whether the question for a specific trial was answered correctly.

The experiment began with two practice passages and comprehension questions, which were immediately followed by the 24 experimental passages and the 24 filler passages presented in a randomized order. Participants read at their own pace, but before moving on from a passage, participants were asked to stare at a row of X's at the bottom of the screen to prevent any extraneous fixations after the passage was finished.

### **Results**

For each of the passages, the target region was the role-filler who performed the scripturally appropriate or inappropriate action (e.g., guitarist), and the post-target region was the following two words (e.g., who played). The data consisted of first pass reading times (the sum of all fixations in a region from first entering the region until leaving the region) for the target and post-target region; second pass reading times (the sum of all fixations in a region following the initial first pass time, including zero times when a region is not refixated) for the target and post-target region; and the percentage of regressions made to the target word from the post-target region. Trials in which the target word was initially skipped, as well as ones in which first and second pass reading times were greater than three standard deviations from the participant's mean, were excluded from analysis. This eliminated approximately 10% of the data. All analyses were considered significant at the standard alpha level of 0.05 and all post-hoc tests were conducted using Fisher's Least Significant Difference.

Table 1 presents the means and standard deviations of the first and second pass reading times in the target and post-target region as well as the percentage of regressions for each condition. One-way ANOVAs were conducted for each reading time measure. No significant difference was found for the first pass reading times for either the target or post-target regions (both  $F$ 's < 1; see Figures 1 and 3, respectively).

Significant differences between the conditions were detected for second pass reading times for the target word,  $F(2,28) = 4.53$ ,  $MSE = 1266.11$ . Fisher's LSD tests indicated that participants in the inappropriate neutral (IN) condition had significantly longer reading times compared to participants in both the inappropriate justified (IJ;  $p = .04$ ) and appropriate neutral (AN;  $p = .007$ ) conditions. However, no significant difference was found in the reading times of participants in the IJ compared to the AN conditions, although participants in the IJ condition had faster reading times ( $p = .40$ ; see Figure 2).

Second pass reading times in the post-target region was also significantly different between conditions,  $F(2,28) = 4.51$ ,  $MSE = 1874.76$ . Participants in the IN condition had significantly longer second pass times on the post-target region compared to either IJ or AN participants ( $p = .01$  and  $.02$ , respectively), which were not significantly different from each other ( $p = .82$ ; see Figure 4).

Finally, the percentage of regressions from the post-target region to the target region was also significant,  $F(2,28) = 3.95$ ,  $MSE = 18.93$ . Participants in the IN condition made significantly more regressions compared to participants in either the IJ ( $p = .02$ ) or AN ( $p = .03$ ) condition. There was no significant difference

between participants in the IJ compared to the AN conditions ( $p = 0.92$ ; see Figure 5).

### **Discussion**

In this experiment, participants read passages about scripted events in which an actor performs a role that was either appropriate or inappropriate based on the script (world knowledge). When the action was inappropriate, it was preceded by context that either justified the script violation or was neutral with regard to the scripted action. This context was then backgrounded, and eye movements were recorded on the target word (the actor) and the prepositional phrase that followed. The results were not significant except for the second pass reading times on the target word (slower in the inappropriate neutral condition) and post-target word (slower in the inappropriate neutral condition), as well as the percentage of regressions (more frequent for the inappropriate neutral condition). These null results are not surprising in light of a critical methodological error I committed.

The error pertains to how participants were assigned to the experimental conditions. Eye-movement effect sizes are generally small, and to detect these effects most studies rely on a within-subjects design, which dramatically increases statistical power by decreasing subject-to-subject variability. For example, in Cook and Myers's study, each subject read passages in all of the experimental conditions, and therefore served as their own "controls". In the present study, participants were mistakenly assigned to a single condition, creating a between-subjects design with only about 10 subjects per condition. As a result, these null results are not

particularly informative<sup>1</sup>. Nevertheless, for the sake of thinking through potential results, I will still speak to trends in the data that might have been significant with a within-subjects design. Therefore, It is important to note that the following discussion is highly speculative, and not backed by data.

*First pass times on the target word*

The first pass results on the target word revealed evidence of greater processing difficulty when the role filler was inappropriate, regardless of whether or not the action had been previously justified ( $M_{IN} = 224$ ;  $M_{IJ} = 227$ ), compared to the appropriate control condition ( $M_{AN} = 211$ ). This trend cannot be explained by Hess et al.'s (1995) lexical reinterpretation hypothesis, which would predict that the presence of a contextual justification should have been sufficient to eliminate any processing difficulties on the target word. However, this trend is consistent with Cook and Myers's (2004) two-stage model. According to their model, when the inappropriate role filler was encountered, a signal is sent to long-term memory that reactivates related information from episodic and semantic memory. In their experiment, the contextual justification was presented immediately before the scripted action, and thus was presumably still accessible in working memory when the inappropriate role filler was encountered. Thus, the initial link was quickly established in the model's first stage with this contextual information, which eliminated any initial processing difficulties due to the inappropriateness of the role filler in their equivalent of the IJ condition. However, the backgrounding section of

---

<sup>1</sup> A within-subjects version of this experiment has been programmed, and data collection is underway as of May 2013. Interested parties may contact the author at [micah.mumper@gmail.com](mailto:micah.mumper@gmail.com) for the updated results.

the present experiment reduces the saliency of the episodic information, and likely requires the episodic information must be reactivated along with world knowledge when the target word is encountered. Thus, in the early processing stage, the interaction between these two sources of information would lead to processing difficulties in the first stage for the IJ condition that was missing from Cook and Myers results. This is exactly the pattern found in the present experiment.

Processing difficulties, which manifest as longer fixation times, were found for both the IN and IJ conditions compared to the AN condition. Thus, the evidence from first-pass fixation times favors the two-stage model.

*First pass times in the post-target region*

The pattern of results found for the first-pass times in the post-target region are more difficult to explain. Cook and Myers (2004) found longer first pass times in the post-target region for both inappropriate conditions ( $M_{IN} = 525$ ;  $M_{IJ} = 500$ ) compared to the appropriate conditions ( $M_{AN} = 476$ ;  $M_{AJ} = 462$ ). They explained this effect as indicating that when subjects read IN texts, they moved on before integration was completed. More interestingly, when their participants read the IJ texts, they also appeared to slow down in the post-target region due to incomplete integration of the target word. Thus, while first-pass times on the target word failed to find evidence of processing difficulties due to the appropriateness of the role-filler, the first pass times on the post-target region did find this effect. This was taken as evidence of stage-two processing, as the initial links in the first stage were constructed with the episodic information, and then evaluated against world knowledge in the second stage. However, the present experiment found evidence of

faster, not slower, reading times in the inappropriate neutral condition compared to the other two conditions ( $M_{IN} = 234$ ;  $M_{IJ} = 249$ ;  $M_{AN} = 249$ ). The reason for this facilitation in the IN condition is unclear, and would not have been predicted by either theory. The facilitation could not be due to a simple explanation that the participants in the IN condition failed to notice the conflict with general world knowledge, as slowdowns were observed in the first pass times for the target word. I will return to this issue.

*Second pass times in, and regressions into, the target region*

The second pass reading times on the target word revealed significantly longer fixations in the IN condition ( $M_{IN} = 97$ ) compared to the IJ ( $M_{IJ} = 64$ ) and AN condition ( $M_{AN} = 50$ ), which did not significantly differ. However, significantly slower second pass times in the IN condition compared to the other two conditions would be predicted by both theories. A difference between the IJ and AN condition would have been more diagnostic, as this difference could not be explained by the lexical reinterpretation hypothesis. According to that theory, if the contextual information was sufficient to override the influence of general world knowledge, then the presence of an episodic justification should eliminate any increased difficulty integrating the target word compared to subjects in the AN condition. However, increased processing difficulty (as shown by longer second pass reading times on the target word) would be indicative of stage two processing in Cook and Myers's (2004) model, as the individual in the IN and IJ condition attempts to integrate the inappropriate role filler.



The percentage of regressions from the post-target area revealed significantly more regressions in the IN condition ( $M_{IN} = 14$ ) compared to the IJ ( $M_{IJ} = 9$ ) and AN condition ( $M_{AN} = 9$ ), which did not significantly differ. The pattern of this measure was similar to the second-pass times, in that IN participants made significantly more regressions compared to IJ or AN participants, however, there was essentially no difference between participants in the IJ and AN condition. This pattern of results fits nicely with the second pass reading times, as the higher percentage of regressions back to the target word would increase the second pass reading times of the target word. This result, however, is difficult to square away with the two-stage model, which would predict that if both the IN and IJ participants were having difficulty integrating the inappropriate role filler, then both of these participants would have made more regressions than the AN participants, in a similar manner to the second-pass times. What this suggests instead is that IJ participants were either a) able to integrate the inappropriate role filler into their situation models before moving on in the text, or, b) when they did regress back to the target word due to incomplete integration, they spent a more time looking at the target word compared to AN participants, but less than IN participants. This pattern makes it clear that the IJ participants had difficulty integrating the inappropriate role filler, but were able to do it faster than the IN participants. Thus, when these results are considered in combination with the other pattern of results, it becomes clear that only the two-stage model of Cook and Myers (2004) could account for this pattern of results.

These results also bear upon the peculiar results of the post-target first pass times. Taken together, the results of the second pass reading time and the

percentage of regressions suggests that participants in the IN condition made significantly more regressions back to the target word from the post-target region, and spent more time looking at the target word after regressing back to it, which is indicative of processing difficulties. Thus, the first pass times in the post-target region may have been shorter because the high number of regressions these participants were making shortened the amount of time the participants initially spent looking at the post-target region. Indeed, an exploratory analysis of the second pass times in the post-target region (a measure not computed by Cook and Myers) revealed that IN participants spent significantly longer in this region ( $M_{IN} = 126$ ) than either the IJ ( $M_{IJ} = 74$ ) or AN ( $M_{AN} = 78$ ) participants, who did not significantly differ. Thus, while the first pass times on the post-target region might have suggested the anomalous result that participants in the IN condition were having less difficulty than the other two conditions, this may not be the case. Instead, the increased number of regressions, and longer second pass times in both the target and post-target region suggest that these participants did actually experience more processing difficulties than the other two conditions, which is what would be expected.

### *Extensions*

As mentioned previously, a revision of this experiment is underway to amend the methodological issue. The results of this experiment provide some initial evidence in favor of the two-stage model, but this requires replication using an appropriate experimental design.

If the results of the corrected study favor the model of Cook and Myers (2004), follow up studies may wish to manipulate resonance as another test of the model's validity. If the first stage is dominated by a resonance type process, the presence of a contextual cue after the backgrounding section that reactivates the contextual justification should be sufficient to replicate the original effects described by Cook and Myers (2004).

Finally, further research on the two stage model needs to clarify how long information may continue to be processed in stage two before the system moves forward in order to more accurately predict post-target effects. It is conceivable that certain pieces of information are processed in stage two longer than others due to episodic manipulations that requires further attention to integrate the information. In the present experiment, participants in the IJ condition showed post-target processing difficulties in the second pass reading time on the target word, whereas, participants in the IN conditioned showed post-target processing difficulties in both the second pass reading times on the target word as well as the percentage of regressions. Understanding how participants handle inconsistent information after moving past the target word is critical to understanding the type of processing that is involved with stage two.

The results of Duffy and Kier (2004) still require a sufficient explanation. The present two-stage model presented here could not predict the results of that study. Recall that they found disambiguating the gender of a character in a stereotypical profession was sufficient to eliminate any observable processing difficulties when the participants encountered the stereotype-violating character later on in the text.

However, several key methodological differences may account for their results. For one, stereotype information may be processed differently than scripts. Future research may wish to vary the type of world knowledge used in the experimental passages (e.g., scripts, stereotypes, historical events). Furthermore, in the present experiment, the episodic justification was written to create an expectation that some role may be violated, but this violation did not occur until the target sentence. In Duffy and Kier's (2004) experiment, the inappropriate role-filler was explicitly disambiguated before the target sentence was encountered. Thus, there may be important differences in how readers integrate explicit contextual information rather than implied. Follow-up studies may wish to see how processing is affected if the role-filler is explicitly disambiguated earlier in the story (i.e., explicitly state that the manager was a member of the rock-band itself). In this case, the strong contextual information may be enough to override the influences of general world knowledge, leading to results similar to what Duffy and Kier found. Indeed, Myers et al. (2000) found that elaborating the target word by increasing the number of mentions eliminated a general world knowledge effect.

### *General Conclusions*

In this experiment, increased processing difficulty was detected after backgrounding the contextual justification in the first-pass times of the IN and IJ conditions. While the results of the present experiment do not allow definitive conclusions to be drawn, this was the critical result needed to support the model proposed by Cook and Myers (2004). The lexical reinterpretation hypothesis could

not have predicted the longer first pass times on the target word for participants in the IJ condition.

In addition, measures of post-target processing difficulties were also found for IN and IJ participants, although the pattern of results was more difficult to interpret. As expected, participants in the IN condition showed the clearest evidence of post-target processing difficulties, indicated by the increased percentage of regressions and longer second pass time on the target word compared to the control AN condition. However, IN participants paradoxically also had faster first pass times on the post-target region, but these were explained as being an artifact of the high percentage of regressions from the post-target region to the target region. Participants in the IJ condition showed evidence of intermediate processing difficulties, although none of the differences from the AN condition were significant. Nevertheless, the trends suggested that IJ participants had longer second pass reading times compared to control participants in the AN condition, but were nearly identical to in terms of percentage of regressions and first pass times in the post-target region. Thus, the data appears to support the two-stage model of Cook and Myers (2004).

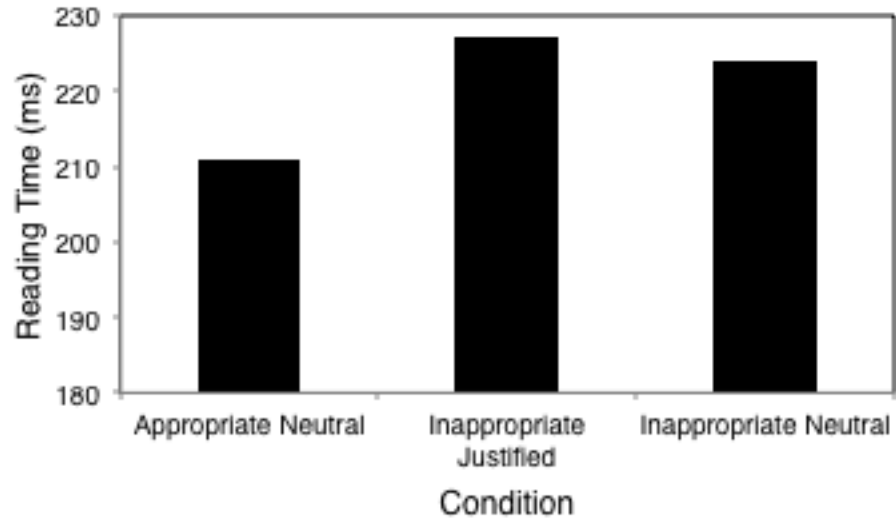
The two-stage model is superior to current theories for several reasons. By assuming early processing can be influenced by either semantic or episodic information, this model can account for the results of studies that find an early influence of either source of memory. Indeed, it is more parsimonious to this author that processing would depend on the saliency of information rather than the particular source. Furthermore, the two-stage model provides a mechanism that can

describe both passive and strategic reactivation of episodic and semantic information. Cook (2005) makes the preliminary case that stage two processing may involve more active processes on the part of the reader in addition to the automatic component. Future research should make use of eye tracking and other on-line methodologies to capture the time-course of processing to bear upon this issue.

Table 1

*Means (standard deviations) first and second pass times (in ms) and percentage of regressions into target region from post-target region.*

	Appropriate Neutral	Inappropriate Justified	Inappropriate Neutral
<i>Target Region</i>			
First Pass	211 (32)	227 (46)	224 (42)
Second Pass	50 (30)	64 (39)	97 (36)
<i>Post-Target Region</i>			
First Pass	249 (80)	249 (50)	234 (49)
Second Pass	78 (45)	74 (40)	126 (45)
<i>Regressions</i>	9 (3)	9 (6)	14 (6)



*Figure 1.* First pass reading times (in ms) for the target region.



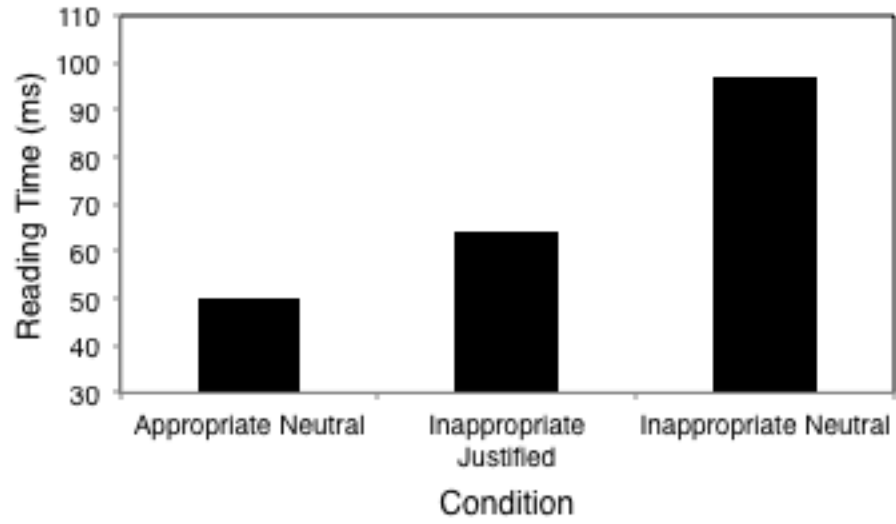


Figure 2. Second pass reading times (in ms) in the target region.

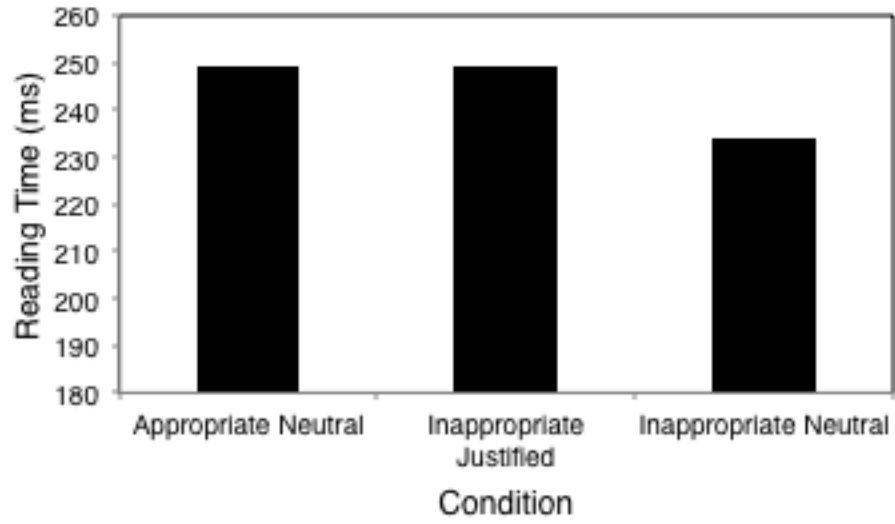


Figure 3. First pass reading times (in ms) for the post-target region.

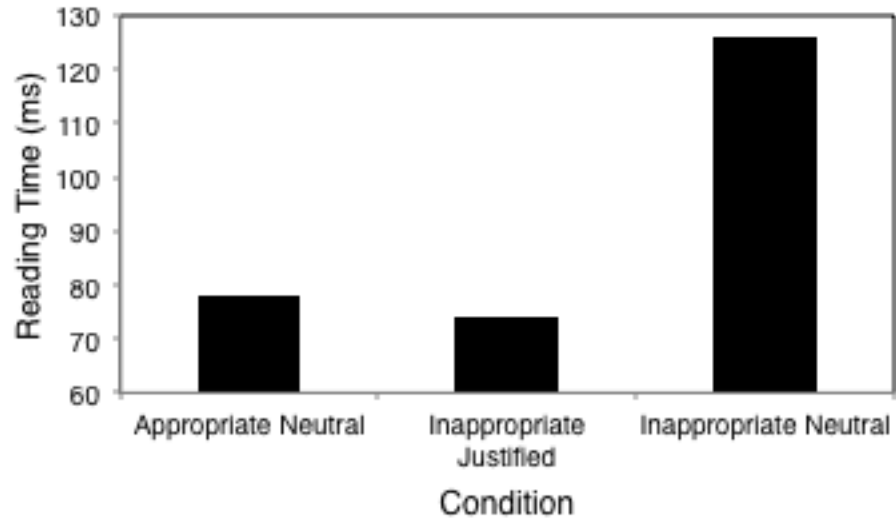
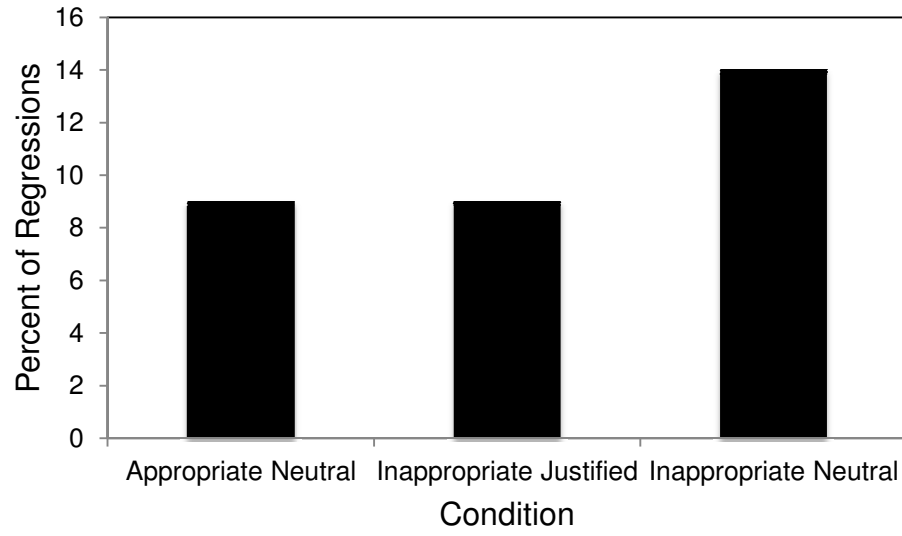


Figure 4. Second pass times (in ms) in the post-target region.



*Figure 5.* Percentage of regressions from post-target region to target-region.

## References

- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1061–1070.
- Cook, A. E. (2005). What have we been missing? The role of general world knowledge in discourse processing. *Discourse Processes, 39*, 265-278.
- Cook, A. E., Halleran, J. G., & O'Brien, E. J. (1998). What is readily available? A memory-based view of text processing. *Discourse Processes, 26*, 109–129.
- Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language, 50*, 268-288.
- Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551-559.
- Foss, D. J. (1982). A discourse on semantic priming. *Cognitive Psychology, 14*, 590–607.
- Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language, 42*, 526–544.
- Gueraud, S. (2005). Updating situation models: The memory-based contribution. *Discourse Processes, 39*, 243-263.
- Hakala, C. M., & O'Brien, E. J. (1995). Strategies for resolving coherence breaks in reading. *Discourse Processes, 19*, 167–185.

- Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, *124*, 62–82.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182.
- Myers, J. L., Cook, A. E., Kambe, G., Mason, R.A., & O'Brien, E. J. (2000). Semantic and episodic effects on bridging inferences. *Discourse Processes*, *29*, 179-199.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*, 131–157.
- O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1200–1210.
- Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*, 527-544.
- Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science*, *14*, 328–333.
- Traxler, M. J., & Foss, D. J. (2000). Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1266–1282.

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory & Cognition, 28*, 1011–1021.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin, 123*, 162-185.

## Appendix A

### Sample Passage [Split by Section]

#### [Introduction]

Lucy was the manager of a small apartment complex.

#### [Neutral Context]

Although she didn't live there and was rarely around to check on things, the residents all felt that the building was in good shape.

#### [Justified Context]

Although she didn't live there, several of her staff members were residents in the complex and she trusted them to keep the place running.

#### [Background]

The residents were a mix of young and old people, families and single people. Nevertheless, they all got along very well, and often did things together. People carpooled to work and often had parties for the entire complex.

#### [Appropriate Role-Filler (Target word in italics, post-target region underlined)]

On the first day of every month, the rent had to be paid by each *tenant* who lived in the building.

#### [Inappropriate Role-Filler (Target word in italics, post-target region underlined)]

On the first day of every month, the rent was collected by a *tenant* who lived in the building.

#### [Closing (Appropriate Condition)]

Most of the tenants had their rent checks ready on time, but occasionally some of the checks did bounce.



[Closing (Inappropriate Condition)]

This tenant was one of Lucy's most trusted staff members.

[Comprehension Question]

Did Lucy live in the apartment complex? {No}