Why do You Read?
Toward a More Comprehensive Model of Reading Comprehension: The Role of Standards of Coherence, Reading Goals, and Interest

by

Regina Christine Calloway

B.S., University of Maryland, College Park, 2013

M.S., University of Pittsburgh, 2016

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Regina Christine Calloway

It was defended on

March 29, 2019

and approved by

Aidan Wright, Assistant Professor, Department of Psychology

Byeong-Young Cho, Associate Professor, Department of Instruction and Learning

Natasha Tokowicz, Associate Professor, Departments of Psychology and Linguistics

Dissertation Director: Charles Perfetti, Distinguished University Professor, Department of Psychology

Why do You Read?
Toward a More Comprehensive Model of Reading Comprehension: The Role of Standards of Coherence, Reading Goals, and Interest

Regina C. Calloway, PhD

University of Pittsburgh, 2019

Readers read for different purposes and the texts they read vary in topic and difficulty. These situational factors influence standards of coherence—how much understanding a reader aims to have for a given text. Three studies examined whether individual differences in reader-based standards of coherence influenced off-line and on-line comprehension. Study 1 designed and evaluated a self-report measure of reader-based standards of coherence. For an adult community sample, an exploratory factor analysis found that the reader-based standards of coherence measure had four factors: 1) intrinsic reading goals, 2) extrinsic reading goals and learning strategies, 3) desire to understand and reading regulation strategies, and 4) desired reading difficulty. The measure predicted readers' reading habits. Study 2 positioned reader-based standards of coherence within a structural equation model of reading comprehension and the findings supported predictions from the Simple View of Reading (Gough & Tunmer, 1986) and the Reading Systems Framework (Perfetti, 1999; Perfetti & Stafura, 2014). College students' listening comprehension and vocabulary knowledge directly affected reading comprehension and decoding ability and reading experience indirectly affected reading comprehension via vocabulary knowledge. Crucially, the structural equation model showed that students with higher reader-based standards of coherence sought out more reading experiences, indirectly affecting vocabulary knowledge. Study 3 tested effects of reader-based standards of coherence, comprehension goal (answering open-ended questions vs. phrase matching), and interest on on-line reading and listening comprehension. Participants with the goal to answer open-ended questions read more slowly than those who completed a phrase matching task, indicating that comprehension goals influenced reading regulation strategies. Additionally, participants with more reading experience and more interest read passages more quickly. Participants across both comprehension goal conditions showed evidence of activating bridging inferences during reading and listening comprehension tasks; however, only participants with high interest showed evidence of activating predictive inferences during reading. Finally, reader-based standards of coherence predicted participants' interest in the passages they comprehended only in more difficult comprehension situations. Overall, the studies demonstrate that reader-based standards of coherence, interest in text material, and reading-related skills help explain sources of comprehension failures in adult readers.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

People read different texts for a variety of reasons. Across situations, readers adopt different standards of coherence, i.e., criteria for understanding (van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011; van den Broek, Lorch, Linderholm, & Gustafson, 2001; van den Broek, Risden, & Husebye-Hartman, 1995). Standards of coherence are influencd by reading goals, text topic, text genre, and readers' ability to implement appropriate strategies across different reading situations (e.g., reading a difficult text more slowly than an easy text). Thus, standards of coherence should influence text representations, including inference generation. Situational factors are not the only ones that influence a reader's level of text understanding. Reader factors, such as enjoyment of reading and comprehension ability, also play a role in text understanding. The present set of studies focused on reader and situational factors that influence text comprehension, with the proposal that there are two types of standards of coherence: 1) situation-based standards of coherence that change across texts and reading goals, and 2) reader-based standards of coherence, which are inherent to individual readers, i.e., given the same reading goal and text, different readers will set different critera for understanding. The present set of studies aims to create a measure of reader-based standards of coherence, position it within a structural equation model of reading comprehension, and test the influence of reader- and situation-based standards of coherence on on-line inference generation.

1.1 Text Representations

Prior research has demonstrated that situation-based standards of coherence, influenced by reading goal, topic, or text expectations, affect readers' mental text representations (e.g., van den Broek et al., 2001; Zwaan, 1994). In their construction-integration model, van Dijk and Kintsch (1983) proposed that readers construct three levels of text representation—surface, propositional, and situational. The surface level includes verbatim features of the text, including individual words and their meanings along with sentences' grammatical structures. The proposition level representation contains meanings of clauses or sentences constructed from the surface level. The final level is the situation level representation, or situation model. The situation model is a referentially rich representation of text meanings, containing activations of relevant background knowledge and inferences combined with the propositional text representation.

Take the following sentences for example, 1) *While Cathy was riding her bike in the park, dark clouds began to gather, and it started to storm. The rain ruined her beautiful sweater* (Perfetti & Stafura, 2014). The surface level constitutes the words within the two sentences. The propositional level includes the meaning representations of the clauses (i.e., Cathy on bike, in park, dark clouds, it stormed, etc.). The situation level includes a referential link between "storm" and "rain". Of importance, readers must have the knowledge that storms can involve rain to infer this connection.

Inferences are a core feature of a reader's situation model because they help readers understand a text by filling in key information such as causal gaps and character goals (Long, Golding, & Graesser, 1992). Factors such as text difficulty, reader goals, reader background knowledge, and reading comprehension ability affect inference generation. These inferences may be generated during encoding (i.e., on-line as a passage is read) or during retrieval (i.e., off-line as a passage is recalled). Readers make local and global inferences, creating links between words, clauses, sentences, paragraphs, and larger stretches of text (Kendeou, van den Broek, Helder, & Karlsson, 2014). If a reader generates strong inferences, they are encoded into memory and a reader's situation model (Keenan, Baillet, & Brown, 1984; Seifert, Robertson, & Black, 1985). When readers do not make inferences that establish coherence among propositions, they may have a reduced understanding of the text (Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006).

1.1.1 Taxonomies of Inferences

Researchers have proposed taxonomies of inferences (e.g., Chikalanga, 1992; Graesser, Singer, & Trabasso, 1994), but inferences are generally classified into two main categories. These categories have been labeled variously as necessary vs. elaborative (Graesser et al., 1994), automatic vs. strategic/effortful (McKoon & Ratcliff, 1992), linguistically constrained vs. reader-constructed (Perfetti & Stafura, 2015), and passive vs. reader-initiated (van den Broek & Helder, 2017). The overarching distinction made by these categories is that some inferences are automatic and necessary for text coherence. Other inferences are unnecessary for text comprehension and are generated when readers are highly engaged. Two types of inferences that have been extensively studied and fit into broad categories of coherence maintaining and elaborative are bridging and predictive inferences.

Bridging inferences are often necessary for text coherence. In the present study, bridging inferences require readers to generate links between text-based propositions across sentences or clauses. For example: 2)*…the husband threw the delicate porcelain vase against the wall. It cost him well over one hundred dollars to replace the vase.* (Potts, Keenan, & Golding, 1988). In the example, for one to integrate the sentences into a single representation of the text's meaning, a proposition resembling “break” must be generated. These bridging inferences are sometimes referred to as backward inferences (e.g., Singer & Ferreira, 1983) and include causal inferences (Kuperberg et al., 2006; Mason & Just, 2004). Sentences that require bridging inferences have longer reading times, which suggest that readers require more time and processing effort to make such inferences (Keenan, Baillet, & Brown, 1984; Experiment 1).

Predictive inferences are sometimes referred to as forward (e.g., Singer & Ferreira, 1983) and elaborative inferences (e.g., Potts et al., 1988). Predictive inferences enrich a reader’s situation model, but are unnecessary for text comprehension (McKoon & Ratcliff, 1986). In example 2, one could make a predictive inference that the vase broke after reading the first sentence about throwing the vase against the wall. However, making the inference prior to reading the second sentence is not necessary.

Both bridging and predictive inferences can be assessed via on-line and off-line measures. On-line measures assess comprehension processing while a word is being read, or immediately after word presentation while the meaning is still active. One example is a probe recognition task, which is useful for capturing reading comprehension and inference generation processes as they unfold (McKoon & Ratcliff, 2015). In a probe recognition task, readers are presented with texts and given a probe word that either relates to an inference, e.g., *break* from example 2, or is unrelated to the text. Readers indicate whether the probe occurred in the previous text. If a reader generates an inference, then the word or proposition should still be active when the probe is shown. Confusion as to whether the probe occurred in the previous text inhibits the decision, producing longer reaction times (RTs) and increased errors to inference-related probes compared to unrelated probes (e.g., Fincher-Kiefer, 1995; McKoon & Ratcliff, 1986, 2017). This task has provided evidence for bridging (e.g. Fincher-Kiefer, 1995) and predictive inferences (e.g., Fincher-Kiefer, 1995; McKoon & Ratcliff, 1986). Paradigms using phrases (Singer & Spear, 2015) or sentences (Zwaan, 1994) as target probes rather than individual words found similar results.

Two complementary bottom-up processes, spreading activation and memory-based resonance, may result in longer RTs for words related to bridging or predictive inferences in probe recognition tasks. Spreading activation is a prospective process that occurs when the semantic activation of a currently read word spreads to semantically related words (Anderson, 1983; McNamara, 2005). Memory-based resonance processes are retrospective and occur when semantic activations from a currently read word resonate with information in memory (Myers & O'Brien, 1998). In probe recognition tasks, when readers encounter a probe word that is semantically related to their representation of a text's meaning, the probe word may have increased activation due to spreading activation from the text representation to the probe word. From a memory-based perspective, the probe word may serve as a retrieval cue to information in memory and have facilitated activation. When readers decide whether the probe word occurred, its facilitated activation from spreading activation and/or memory-based resonance should result in confusion as to whether the word actually occurred. This confusion ultimately results in longer response times and more inaccurate responses. The inhibition effect caused by confusion is in contrast to naming and lexical decision tasks, i.e., indicating whether the probe is a legal word, in which there is facilitation for activated words resulting in shorter latencies (e.g., Potts et al., 1988). Although many studies use on-line paradigms to assess inference generation, many factors influence whether readers will make inferences.

1.1.2 Influences on Text Representations

In addition to findings that readers make situation-level representations containing inferences, both reader and situation factors should affect the levels of text representations. One situation factor that influences situation-based standards of coherence is a person's reading goal, which can be extrinsic and set by task instructions, or intrinsic and set by the reader (Narvaez, van den Broek, & Ruiz, 1999; van den Broek et al., 2011). With a focus on the influence of extrinsic reading goals on reading comprehension, using a think-aloud paradigm van den Broek et al. (2001) found that readers instructed to read an expository text for study made more elaborative and explanatory (e.g., bridging) inferences, paraphrases, and repetitions than those instructed to read the same text for entertainment. In contrast, participants who read for entertainment made more associations and text evaluations, e.g., "this statement is weird". Readers with a goal to study also spend more time re-reading than those with a goal to read for entertainment (Yeari, van den Broek, & Oudega, 2015). These findings suggest that readers' situation-based standards of coherence, influenced by extrinsic goals, can affect their reading behavior and propositional links within their situation models.

In a study examining the influence of readers' text expectations, Zwaan (1994) instructed college readers that they were reading news stories or literary stories. Because news stories are meant to convey factual information and require referential links, readers should create situation-level text representations. Literary stories may not always convey factual information and readers may read such stories to enjoy a specific plot or genre. Thus, for literary stories, readers may focus more on peripheral events, creating more surface- or proposition-level representations. Of import, the texts readers read were identical, half were actual news stories and half were literary stories, only reader expectations differed. When comparing the two groups of readers, Zwaan found that participants who read the texts as news stories had more situation-level representations and that those who read the texts as literary texts had more surface-level representations. Activation of proposition-level text representations did not differ between the two groups. In a second experiment, Zwaan found that participants who read stories as literary stories remembered unnecessary, elaborative inferences more than those who read the same stories as news stories. Zwaan concluded that readers allocated their attention differentially depending on their text genre expectations.

These previous experiments demonstrate that extrinsic reading goals influence text representations, however intrinsic reading goals and reader engagement are also related to text comprehension and reading persistence. Reading engagement involves positive affect during reading, interest, active reading, using appropriate reading strategies, motivation, and building knowledge through reading (Guthrie et al., 2004). Less-skilled readers report lower enjoyment—a component of engagement—during reading (Crossley, Skalicky, Dascalu, McNamara, & Kyle, 2017), which may lower their situation-based standards of coherence threshold and negatively affect comprehension. Interest can be limited to particular texts, some texts are more interesting than others, or specific to individuals, people are interested in different topics. Bernstein (1955) focused on text-specific interest and found that ninth grade students had higher comprehension scores for a more interesting story compared to a less interesting story, even when the two stories were matched on text difficulty. In a study of individual interest, McWhaw and Abrami (2001) found that 11th grade students who read a Psychology text that they rated as interesting identified more main ideas and used more metacognitive strategies for the text compared to those who found the text uninteresting. Thus, giving readers texts that match their topic interests should increase their motivation and engagement during reading (e.g., Belloni & Jongsma, 1978; Fulmer &

Frijters, 2011; McWhaw & Abrami, 2001) and, consequently, their situation-based standards of coherence.

For example, Fulmer and Frijters (2011) found that middle-school children who read a difficult text related to their most-liked topic reported higher enjoyment and perceived competence about reading the text than children who read passages related to topics they rated as least-liked. Additionally, children who read passages on topics of interest were more likely to continue reading the passage, even when it was above their reading level. The results indicate that children's reading engagement can offset negative effects of reading difficult texts and similar patterns have been reported with adults. Nell (1988; Study 5) found that college students rated enjoyable books that they had recently read as requiring less effort to read compared to when they read something enjoyable in a distracting situation, a work book, or an unenjoyable book for pleasure. When readers find texts more interesting, they are more motivated and less likely to be distracted, which leads to increased comprehension (Unsworth & McMillan, 2013).

1.1.3 Reading-related Skills

Various reader factors relate to reading comprehension. Readers may differ in reading-specific skills, such as decoding, or other individual abilities, such as reading experience and intrinsic motivation. If readers have poor decoding skills, they may spend more cognitive resources trying to identify words and access their meanings. This increased effort results in fewer resources remaining for inference generation (Bell & Perfetti, 1994; Gough & Tunmer, 1986).

At a lower-level, phonological awareness and orthographic knowledge relate to decoding ability (Landi, 2010; Shankweiler et al., 1999), and good readers have strong phoneme-grapheme correspondences (Pratt & Brady, 1988; Share, 1995). In a meta-analysis of 110 studies, García and

Cain (2014) found that across all ages (children to adults), decoding ability correlated with reading comprehension. This relationship was moderated by age; older readers had a weaker relationship between decoding ability and reading comprehension than younger readers.

In addition to decoding, the relationship between vocabulary knowledge (i.e., knowledge of word meanings) and reading comprehension has been assessed. Although vocabulary knowledge refers to word meanings, during text comprehension readers must be able to access word meanings via links to their visual word-forms. Thus, in addition to vocabulary knowledge, readers need strong lexical representations that include interactions among phonological, orthographic, and semantic knowledge (Perfetti & Hart, 2002; Perfetti, 2007). For instance, the visual word-form *bike* should activate its phonological and semantic representations, which strengthen the orthographic representation. Their combined activation facilitates word identification at both word-form and semantic levels of representation.

Listening comprehension, another predictor of reading comprehension, also requires vocabulary knowledge, but it does not require written word decoding. Listeners have cues that are unavailable during reading, such as intonation and prosody. Because of these additional cues, and eliminating the need for written word decoding, listening comprehension abilities should be higher than reading comprehension abilities. Braze, Tabor, Shankweiler, and Mencl (2007) conducted a study measuring listening and reading comprehension in young adults and, using a median split, found that less-skilled readers performed higher on listening comprehension assessments than reading comprehension assessments. More-skilled readers, however, showed no difference between performance on listening and reading comprehension assessments. Braze et al. proposed that less-skilled readers struggled with reading-specific skills and required more vocabulary knowledge to help them with word identification. This conclusion aligns with previous findings that less-skilled readers rely on context information more than more-skilled readers (Perfetti, Goldman, & Hogaboam, 1979). An alternative explanation for the difference between reading and listening comprehension for less-skilled readers is that less-skilled readers had higher engagement during the listening task than the reading task. The present study tests differences between listening and reading comprehension while controlling for engagement in each modality (Study 3).

Of note, for inference generation to occur, comprehenders must be able to engage in comprehension monitoring, the ability to evaluate and regulate understanding (Baker, 1985). Thus, when readers detect a comprehension failure, they must employ the correct strategy to revise the failure (Wagoner, 1983). One widely used measure of comprehension monitoring is inconsistency detection. When readers detect an inconsistency, they read more-slowly and re-read material (Albrecht & O'Brien, 1993; Kim, Vorstius, & Radach, 2018; O'Brien, Rizzella, Albrecht, & Halleran, 1998). In a study with sixth grade students, Zabrucky and Ratner (1992) found that both good and poor readers detected inconsistencies, however good readers were better at regulating reading strategies for inconsistencies detected in expository texts. The two groups did not differ in employing reading regulation strategies for narrative texts, suggesting that text difficulty (e.g., narrative vs. expository texts) may influence one's ability to use appropriate regulation strategies (Zabrucky & Ratner, 1992).

Other research demonstrates that skilled and less-skilled readers differ in whether they use background knowledge to make inferences (Cain, Oakhill, Barnes, & Bryant, 2001; Murray & Burke, 2003; Long, Oppy, and Seely, 1994; Ng, Payne, Steen, Stine-Morrow, & Federmeier, 2017; Perfetti, Yang, & Schmalhofer, 2008). Cain et al. (2001) assessed reading comprehension with seven- and eight-year-olds after giving all children the same background information about an alien planet, which was the main topic of the stories they read. Skilled readers outperformed less-skilled readers on off-line open-ended questions assessing literal comprehension, bridging inferences, and elaborative inferences (e.g., additional details unnecessary for text comprehension). Although the less-skilled readers could recall necessary background information when asked, they often failed to use that information to accurately answer questions to make bridging and elaborative inferences. Thus, children must be able to activate relevant background knowledge to successfully answer inference-related questions (Cain et al., 2001; Carlson et al., 2014). Because readers answered questions off-line, it is unclear whether they would have activated bridging or elaborative inferences without the question prompts. Additional research examining on-line bridging inferences with a naming task found that adult skilled readers had longer naming times for inference-related probes, indicating that they activated meanings related to necessary bridging inferences and experienced inhibition (Long et al., 1994). However, Long et al. found that less-skilled readers did not activate the bridging inferences, indicated by similar naming times for control and bridging-related probe words.

Whether readers generate predictive inferences also varied by reading ability, however findings have been mixed. In a naming study, West and Stanovich (1978) compared naming times for fourth graders, sixth graders, and college students reading single sentences with the final word removed. West and Stanovich found that all students had facilitated naming times for predictable words compared to unpredictable words and words in isolation. Fourth and sixth graders, but not college students, had longer naming times for unpredictable words than words in isolation. Additionally, across all age groups, less-skilled readers benefited more from reading predictable words in context than the same words in isolation, suggesting that they make use of context information during reading. In a similar study with fifth grade students, Perfetti et al. (1979) found that less-skilled readers benefited more from naming words in context (Experiment 2). When participants read short stories and predicted an upcoming word prior to its presentation (i.e., a cloze task), naming times for predicted words were faster than for unpredicted words. Although less-skilled readers were less accurate in predicting specific words, they did not differ from skilled readers in predicting context-appropriate words. Overall, the two studies suggest that less-skilled readers make predictive inferences, and prior context facilitates their word identification, measured via naming times, more than skilled readers.

In contrast, a study by Murray and Burke (2003) found different effects of reading skill on prediction. In Murray and Burke's study, participants read short passages and were prompted to read a probe word that occurred 500 ms after the end of the final sentence. More-skilled readers had faster naming times for words related to predictive inferences than unrelated control words, but moderate- and less-skilled readers did not show this effect (Experiment 1). The absence of differences in naming latencies for low- and moderate-skilled readers may be because lower-skilled readers did not have enough time to generate the inference or the task itself did not encourage predictive inference generation to a high degree. West and Stanovich (1978) and Perfetti et al. (1979) included texts in which the final word of a sentence was absent. Furthermore, Perfetti et al. explicitly instructed participants to predict the final words (Experiment 3). Such conditions may have encouraged predictive inferencing more than the complete passages used in Murray and Burke's study.

Event-related potential (ERP) studies elucidate more specific situations in which skilled and less-skilled readers make predictive inferences. Skilled readers made on-line predictive inferences within strongly and weakly constraining sentences, indicated by a reduced N400—a marker of semantic congruence peaking around 400 ms post-stimulus—for predicted compared to unpredicted words (Ng et al., 2007). Less-skilled readers showed this effect only within strongly constraining sentences (Ng et al.). Less-skilled readers also show this effect at the beginning of a second sentence when the first sentence has a constraining context (e.g., Perfetti, et al., 2008). Perfetti et al. found that more-skilled readers, however, did not show a reduced N400 for predicted words compared to unpredicted words across a sentence boundary, but they did exhibit faster and larger effects of establishing co-reference compared to less-skilled readers. Skilled readers may not have made predictive inferences because sentence beginnings allow for multiple possible continuations and predictions are likely to be inaccurate (Calloway & Perfetti, 2017; Helder, Perfetti, van den Broek, Stafura, & Calloway, 2019) and word predictability is lower at sentence beginnings than endings (Luke & Christiansan, 2016). Thus, more-skilled readers may have used sentence boundaries as cues to inhibit predictive inference generation more than less-skilled readers. Overall, studies of individual differences in reading demonstrate that both text-based factors and reading comprehension skills affect inference generation.

1.2 Reading Comprehension Frameworks

Various reading comprehension frameworks provide proposals of how reading-related skills influence reading comprehension. The Simple View of Reading (SVR; Gough & Tunmer, 1986) highlights the role of listening comprehension and decoding ability. The Reading Systems Framework (Perfetti, 1999; Perfetti & Stafura, 2014) focuses on the role of vocabulary knowledge in reading comprehension. The Landscape Model (Tzeng, van den Broek, Kendeou, & Lee, 2005; van den Broek & Helder, 2017) shifts focus to the importance of standards of coherence on comprehension processes employed during reading. The next section reviews these three frameworks—the Simple View of Reading, the Reading Systems Framework, and the Landscape Model—in more detail.

According to Gough and Tunmer's (1986) SVR, reading comprehension is the arithmetic product of decoding and oral language comprehension and for adult readers these skills are often slow to change (National Research Council, 2012). According to the SVR, oral language comprehension encapsulates the semantic, syntactic, and pragmatic knowledge a reader has about language, and decoding is a reading-specific process. Both decoding and oral language ability are necessary for reading comprehension. If readers lack automatic decoding abilities, then they have fewer resources available for high-level inference making processes (Bell & Perfetti, 1994; Perfetti & Hart, 2002; Stanovich, 1986). If a reader's spoken language ability is lacking, he or she may not have the skills necessary to make inferences. Importantly, this framework posits that semantic knowledge is subsumed by oral language comprehension.

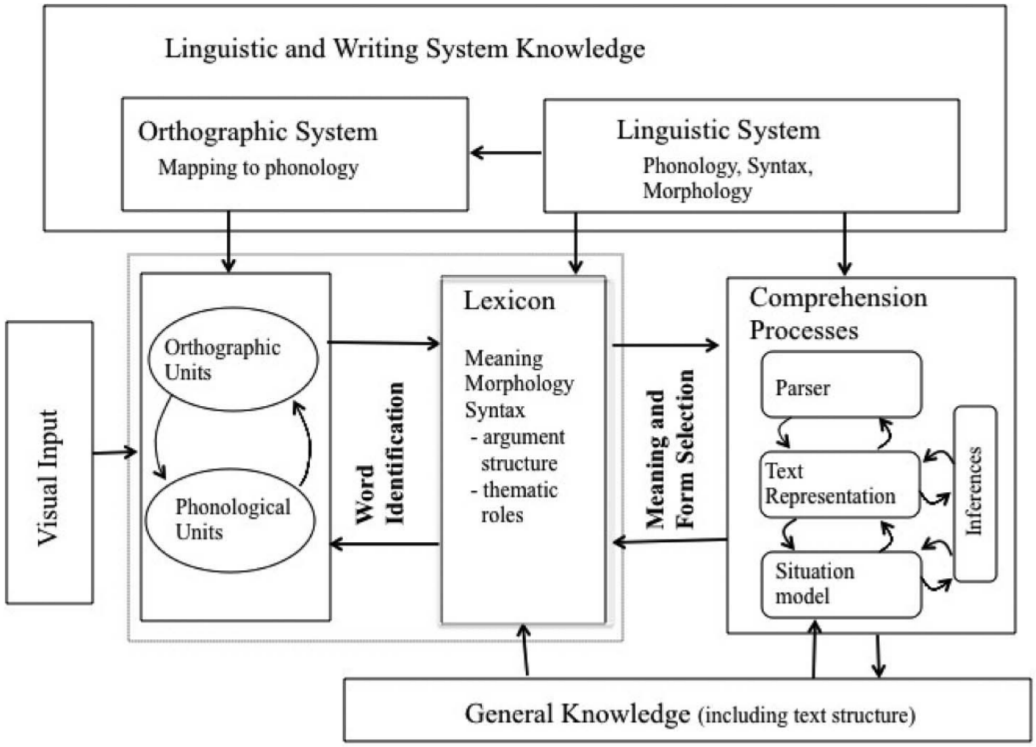

Figure 1. The Reading Systems Framework (Perfetti & Stafura, 2014)

The Reading Systems Framework (Figure 1; Perfetti, 1999; Perfertti & Stafura, 2014) has a more specified approach, defining knowledge sources, comprehension processes, and links between them. Within this framework there are three knowledge sources: orthographic, linguistic, and general knowledge. Reading processes such as decoding, semantic access, meaning integration, and inference processes draw on these three knowledge sources during comprehension and relations among reading processes and knowledge sources are bi-directional. For example, general knowledge may aid inference generation during reading, and information learned during reading informs general knowledge. Similarly, a reader's ability to access word meanings aids both literal and inferential comprehension processes, and comprehension processes may facilitate learning new lexical items. Unlike the SVR, the Reading Systems Framework does not explicitly specify a role for oral language comprehension. Additionally, whereas the SVR positions oral language comprehension as a skill that encompasses semantic, pragmatic, and syntactic knowledge, the Reading Systems Framework separates these into different knowledge sources and reading processes, allowing for more targeted hypotheses about the source of comprehension difficulties. Finally, whereas the SVR specifies a direct link between decoding and reading comprehension, the Reading Systems Framework places lexical knowledge as a link between the two.

Whereas the SVR and Reading Systems Framework focus on reading-related skills, the Landscape Model incorporates situation-based standards of coherence as well. As an extension of the Landscape Model (Tzeng et al., 2005), van den Broek and Helder (2017) proposed a framework that situates reading processes on a continuum. Processes at one end of the continuum require little effort, tend to be automatic, and aid literal comprehension such as semantic access and co-referential binding. Processes at the opposite end require more effort and include inference generation, re-reading, and note taking. Whether readers are engaged in more automatic or effortful processes depend on readers' background knowledge, comprehension ability, and situation-based standards of coherence. Situation-based standards of coherence should interact with reading ability. For example, text genre may lead a more-skilled reader to adopt standards different from a less-skilled reader. A difficult, uninteresting text may lower engagement more for less-skilled readers than more-skilled readers.

Structural equation models (SEM) and other multivariate approaches assessing relationships among decoding, reading comprehension, listening comprehension, and individual differences provide evidence for claims made by each of these frameworks (e.g., Coppens, Tellings, Schreuder, & Verhoeven, 2013; Cutting & Scarborough, 2006; Kendeou, van den Broek, White, & Lynch, 2009; Landi, 2010; Taylor & Perfetti, 2016). In many multivariate analyses that test assumptions made by the SVR, vocabulary knowledge is often assessed orally (e.g., Braze et al., 2007; Braze et al., 2016; Tunmer & Chapman; 2012; Wise, Sevcik, Morris, Lovett, & Wolf, 2007).

Using hierarchical regressions to test the assumptions of the SVR, Braze et al. (2007), tested whether vocabulary knowledge (assessed via oral receptive and expressive measures) accounted for variance in reading comprehension in young adults after accounting for decoding ability and listening comprehension. Braze et al. (2007) found that all three skills, decoding, listening comprehension, and vocabulary knowledge, predicted reading comprehension, suggesting that vocabulary knowledge has an independent influence on reading comprehension.

Using a similar analysis with a fifth-grade sample, Tunmer and Chapman (2012) found that vocabulary knowledge (assessed by an oral receptive measure) was a significant predictor of reading comprehension when listening comprehension, word recognition, and sound-letter knowledge were accounted for. In a SEM with the same sample, decoding and listening comprehension latent factors predicted reading comprehension. However, vocabulary indicators significantly loaded onto a listening comprehension latent factor, suggesting that vocabulary knowledge does not have an independent contribution to reading comprehension. In a SEM study with adults, Braze et al. (2016) also tested the SVR and assessed whether vocabulary knowledge (assessed via oral expressive and receptive measures) served as an independent predictor of reading comprehension. Braze et al. found that vocabulary knowledge measures loaded onto a listening comprehension latent factor, rather than forming a separate latent factor.

Overall, the hierarchical regression analyses and SEMs provided different results. Hierarchical regression analyses suggested that vocabulary knowledge was an independent predictor of reading comprehension; however, the SEMs provided evidence for the SVR, with decoding and listening comprehension predicting reading comprehension. In SEMs, vocabulary was a component of listening comprehension. Braze et al. (2016) proposed that oral vocabulary measures may assess relevant knowledge necessary for listening comprehension not fully encompassed by the measures used to assess listening comprehension. Thus, in the regression analysis, vocabulary knowledge served as a unique predictor of reading comprehension, however in SEMs, it was a component of listening comprehension.

Although SEM results appear to support the SVR (e.g., Braze et al., 2016; Tunmer & Chapman, 2012), both studies used oral vocabulary measures. Because reading is a visual process, linking orthographic word-form representations to semantic representations is a necessary part of reading (Perfetti & Hart, 2002; Perfetti, 2007) and receptive vocabulary from written language is also important to assess. Vocabulary knowledge assessments that require readers to use visual word-forms to access semantic information may provide a unique contribution to reading comprehension. Thus, the present study assesses both oral and visual vocabulary knowledge.

Vocabulary knowledge is correlated with reading comprehension (e.g., Landi, 2010; Perfetti & Hart, 2001; Shankweiler, Dreyer, Lundquist, & Dickinson, 1996). Skilled reading involves accessing word meanings and combining them to create a mental representation. In addition to the vocabulary-comprehension relationship, decoding and vocabulary are significantly correlated (e.g., Braze et al., 2007; Landi, 2010; Shankweiler et al., 1996). The relationship between decoding and vocabulary knowledge and between vocabulary knowledge and comprehension suggest that vocabulary plays a crucial role in linking decoding and comprehension (Perfetti & Hart, 2002; Perfetti, Landi, & Oakhill, 2005; Stafura & Perfetti, 2014).

Reading experience also influences vocabulary knowledge, word identification, and reading comprehension (Cunningham & Stanovich, 1997; Martin-Chang & Gould, 2008;

Stanovich & Cunningham, 1992). People with more reading experience have stronger lexical representations that support word identification processes. A study by Bolger, Balass, Landi, and Perfetti (2008) found that participants who learned new words across four different contexts learned them better than those who learned words presented four times in the same context. Thus, increased reading experiences across different contexts promotes learning new words. In addition to supported word learning, adults with more reading experience re-read words less frequently, have shorter refixation durations, and skip more words during reading (Taylor & Perfetti, 2016).

Previous studies demonstrate that reading experience results in increased word identification, vocabulary knowledge, and comprehension, but what leads readers to engage in reading activities? Reading motivation, a drive to seek out reading activities is associated with reading behavior (e.g., Wigfield & Guthrie, 1997). In a 12-week intervention study with fourth graders, Wigfield et al. (2008) found that combined motivation strategy and reading strategy instruction resulted in higher reading engagement and reading comprehension than reading strategy instruction alone. Using their Motivation for Reading Questionnaire (MRQ), Wigfield and Guthrie (1997) showed that fourth and fifth grade students with higher intrinsic motivation had increased breadth of reading (i.e., number of texts read) and time spent reading compared to students with lower intrinsic motivation. With the MRQ as a template, Schutte and Malouff (2007) created a reading questionnaire for adults (the Adult Motivation for Reading Scale; AMRS). The AMRS has four categories, reading as part of self, reading efficacy, reading for recognition (from others), and reading to do well in other realms. Increased scores on the AMRS were associated with increased enjoyment during reading, reading frequency, and time spent reading. Thus, reading motivation is associated with increased reading experiences.

Another factor that is related to reading behavior and reading comprehension is need for cognition, which relates to one's enjoyment of thinking about or understanding complex problems and situations (Cohen, Stotland, & Wolfe, 1955). Dai and Wang (2007) found that need for cognition predicted reading comprehension for expository and narrative texts. Juric (2017) found that need for cognition predicted number of books read and reading frequency.

Although reading motivation and need for cognition are related to reading behavior and reading comprehension, I propose that reader-based standards of coherence, a reader's need to understand text, may also predict reading behavior and comprehension. Reader-based standards of coherence are similar to need for cognition and the two constructs may be related, those with a high need for cognition may also have high reader-based standards of coherence. Whereas need for cognition is broad, referring to general problem-solving, reader-based standards of coherence are specific to understanding *texts*. Some readers may, on average, set higher standards than others (e.g., a standard that requires a situation-level representation vs. a standard that requires a surface-level representation). Students with the same reading goal of studying for an exam may internalize that goal differently and adopt different strategies to reach that goal, such as highlighting, re-reading, or note taking. Additionally, some readers may fail to adjust their standards appropriately for a given situation. In summary, each reader sets situation-based standards of coherence that may be influenced by the reader's typical criteria for understanding, i.e., reader-based standards of coherence.

The present project posits that both reader- and situation-based standards of coherence influence text representations. Those with high standards of coherence should construct situation-level representations that include inferences. The present study uses structural equation modeling and an experiment to test the contribution of reader- and situation-based standards of coherence to off-line and on-line inference generation.

1.3 Project Aims

The current project addresses three aims related to the roles of reader- and situation-based standards of coherence in off-line and on-line reading comprehension. A structural equation model focuses on the role of reader-based standards of coherence within a reading comprehension model and an experiment focuses on the dynamic nature of situation-based standards of coherence and its relation to reader-based standards of coherence during on-line comprehension. The **first aim** is to develop a measure of reader-based standards of coherence.

The **second aim** of the project is to test a model of reading comprehension that incorporates relationships among listening comprehension, vocabulary knowledge, decoding, reading comprehension, reading experience, and reader-based standards of coherence. Literal and inference-based comprehension are separated within the model.

The **third aim** of the project is to experimentally test whether readers' situation-based standards of coherence across multiple texts affect inference generation. Whereas the structural equation model tests the interrelationships among off-line reading and spoken language skills and reader-based standards of coherence, the experiment captures the role of both reader- and situation-based standards of coherence during on-line literal and inferential comprehension, as they change across texts and reading goals.

2.0 Study 1: Creating a Measure of Reader-based Standards of Coherence

The aim of Study 1 was to develop a measure of reader-based standards of coherence. Although there is empirical evidence that situation-based standards of coherence influence the number of inferences generated—e.g., extrinsic reading goals influence inferences and the amount of information recalled (van den Broek et al., 2001)—there is no measure of readers' reader-based standards of coherence. To measure reader-based standards of coherence, a questionnaire targeting typical reading goals, desire to understand, and comprehension monitoring was developed and refined to an exemplar stage using an exploratory factor analysis (EFA). Two of these categories, typical reading goals and comprehension monitoring, have been studied to some degree. Although there is one known measure of adult reading motivation (i.e., AMRS; Schutte & Malouff, 2007; for a review see Davis, Tonks, Hock, Wang, & Rodriguez, 2018) and the use of reading strategies has been studied (e.g., Taraban, Rynearson, & Kerr, 2000), the core of the reader-based standards of coherence measure is to assess a reader's desire to understand what is read. A reader's desire to understand what is read should be related to reading goals and comprehension monitoring strategies that will lead to high levels of understanding. A person's desire to understand what is read may also be related to a his or her need for cognition, a broader desire for problem-solving and engaging in complex tasks. Because need for cognition is related to reading habits (e.g., Juric, 2017) the present study tests whether, when accounting for need for cognition, reader-based standards of coherence are still related to readers' reading habits.

2.1 Methods

2.1.1 Participants

The study included 205 participants recruited via Amazon Mechanical Turk (MTurk). Participants received $2.00 as compensation and one participant was excluded because of patterned responses to the survey. Of the remaining 204 participants, 112 were male and 92 were female (no participant selected "Other"). Participants ranged in age from 18 to 71 (M_{age} = 31.96, *SD* = 8.58). Of these participants, 20 had a high school diploma or GED equivalent, 20 had an Associate degree, 53 completed some college, 81 had a Bachelor's degree, and 30 had a Master's or Doctoral degree. See Table 1 for a breakdown of ethnicity.

Table 1. Racial and Ethnic Background of Participants

Ethnicity	*n*	*Percentage*
White	163	79.90%
Black or African American	17	8.30%
Asian	9	4.40%
Hispanic or Latino	11	5.40%
Native American or Alaska Native	2	1%
Native Hawaiian or Pacific Islander	0	0%
Other	2	1%

2.1.2 Measures

Participants completed three measures, the reader-based standards of coherence measure, need for cognition questionnaire (Cacioppo, Petty, & Kao, 1984; Appendix A), and a shorted version of a reading habits questionnaire (Finucci, Isaacs, Whitehouse & Childs, 1982; Finucci, Whitehouse, Isaacs & Childs, 1984; Lefly & Pennington, 2000; Parault & Williams, 2009; Appendix B).

2.1.2.1 Need for Cognition

The need for cognition questionnaire consisted of 18 items. This questionnaire related to preferences for problem solving, including complex and abstract problems. Nine items were reverse scaled. Participants rated whether they agreed that a statement is characteristic of them on a Likert scale ranging from strongly disagree (1) to strongly agree (7). The score range of this test is -72 to 72.

2.1.2.2 Adult Reading History Questionnaire (ARHQ)

As a measure of participants' reading habits, participants also completed a subset of items from the Adult Reading History Questionnaire (ARHQ; Finucci et al.., 1982, 1984; Lefly & Pennington, 2000) and one question from the Reading Activity Questionnaire (Parault & Williams, 2009). The questions targeted readers' reading frequency, time spent reading, and types of reading materials (Appendix B). Scores for this subset of the ARHQ ranged from 0 – 23 (excluding the question about reading materials). The question regarding types of reading materials stemmed from the Reading Activity Questionnaire and consisted of three broad categories: 1) magazines (online and in print), 2) newspapers (online and in print), and 3) books (non-fiction and fiction). The "Books" category included non-fiction books (self-help, autobiography, biography, history, nature, and sports) and fiction books (romance, mystery/adventure, fantasy/science fiction, and literature). The survey also had an "other" category in which participants could type out additional genres they read. Participants could select multiple types of reading materials.

2.1.2.3 Reader-based Standards of Coherence Questionnaire

After deciding on the three categories for the standards of coherence measure (typical reading goals, desire to understand, and comprehension monitoring), exemplar items were created. The measure also included 6 items adapted from the AMRS (Schutte & Mauloff, 2007). Finally, open-ended responses from 11 undergraduate students, graduate students, and post-doctoral researchers on the following three questions were used to create additional items: 1) Why do you generally read?, 2) What motivates you to understand a text?, and 3) What reading strategies do you use to understand a text, and in what situation? The reader-based standards of coherence measure included 87 items across three hypothesized categories: typical reading goals (30 items), desire to understand (30 items), and comprehension monitoring (27 items). Of the 87 items, 24 were reverse scaled. Similar to the need for cognition questionnaire, participants rated whether they agreed that a statement was characteristic of them on a Likert scale ranging from strongly disagree (1) to strongly agree (7). Please see Appendix C for a full list of the 87 items.

2.1.3 Procedure

Participants completed the reader-based standards of coherence measure, the subset of the adult reading history questionnaire, and the need for cognition questionnaire. All questionnaires were administered via Qualtrics and took approximately 15 minutes to complete.

2.1.4 Analysis Procedure

The analyses were carried out in two steps. First, a principal axis factor (PAF) analysis, a type of EFA, was conducted to extract the items that significantly loaded onto the factors identified from a parallel analysis. A parallel analysis estimated the number of factors identified from a randomly sampled set of the data and a randomly generated set of data. The number of factors to retain should be fewer than the number identified by the parallel analysis. A PAF analysis was chosen because it is useful for non-normal data (Bartholomew, 1980; Fabrigar, Wegener, MacCallum, & Strahan; 1999), which was the case in the present study. For the PAF analysis, factor loadings greater than .5 were considered significant. Multiple fit indices assessed model fit. Non-significant χ^2 values indicate good fit. The root mean square error of approximation (RMSEA) and Tucker-Lewis Index (TLI) were also included to assess model fit. TLI values > .90 indicate acceptable fit and values > .95 indicate good model fit (Hu & Bentler, 1999). RMSEA values < .08 and < .05 indicate acceptable and good fit, respectively (Hu & Bentler).

After the first step, the second step used hierarchical regressions with the finalized version of the reader-based standards of coherence measure to determine whether it accounted for variance in people's reading habits (ARHQ) when controlling for need for cognition. Analyses for need for cognition and reader-based standards of coherence were conducted separately, then both were entered into a single model. If the change in variance accounted for is significant when the reader-based standards of coherence measure is entered into the model second, it would indicate that reader-based standards of coherence accounts for additional variance not explained by need for cognition. Furthermore, if there is a non-significant change in variance when need for cognition is entered second, it would suggest that the new reader-based standards of coherence measure is a more fine-grain predictor of reading habits than need for cognition.

2.2 Results

The average score for the ARHQ, which measured readers' reading habits, was 11.39 (*SD* = 3.86). Scores ranged from two to 23. Separate from the reading habits numerical score, participants indicated the reading materials they read. A χ^2 test of independence tested whether there were differences among the types of reading materials that people reported reading. Overall, there was a difference among the various reading materials, $\chi^2(15, N = 204) = 214.16, p < .001$. People reported reading science-fiction/fantasy, newspapers online, and mystery/adventure more than magazines in print, newspapers in print, romance, sports, nature, and self-help books. In comparing online vs. print reading, people read both newspapers, χ^2 (1, $N = 204$) = 40.47, $p < .001$, and magazines, χ^2 (1, $N = 204$) = 6.0, $p = .014$, online more than in print. People read fiction more than non-fiction, χ^2 (1, $N = 204$) = 12.45, $p < .001$. Additional analyses examined whether age, gender, ethnicity, or education level influenced the types of materials read. Overall, gender influenced the types of books that people reported reading, χ^2 (15, $N = 204$) = 52.33, $p < .001$. This was primarily driven by romance and sports reading. Whereas men read more about sports than women, χ^2 (1, $N = 204$) = 19.88, $p < .001$, women read more about romance than men, χ^2 (1, $N = 204$) = 18.79, $p < .001$. There were no differences in types of genres read by age (grouped as: 18-24, 25-29, 30-34, 35-39, and 40+), ethnicity (White and Non-white), or education level (high school, some college, Associate degree, Bachelor's degree, and Master's/Doctoral degree; *p*s > .31).

The average score for need for cognition was 13.55 (*SD* = 23.92) and the measure had good internal consistency ($\alpha = .94$, 95% *CI* = .94 - .97). Scores ranged from -51 to 54. Need for cognition was positively correlated with the score from the ARHQ ($r = .42$, $p < .001$). Before assessing whether the reader-based standards of coherence measure related to need for cognition and readers' reading habits, parallel and PAF analyses estimated the number of factors and items to retain in the reader-based standards of coherence measure.

2.2.1 Principal Axis Factor Analysis on Reader-based Standards of Coherence

The reader-based standards of coherence measure was designed to have three factors; however, it is possible that the data reflect a different factor structure than hypothesized (typical reading goals, desire for understanding, and comprehension monitoring). Thus, a parallel analysis using the lavaan package in R estimated the number of factors to retain (R v. 3.5.0; Rosseel, 2012). Although the parallel analysis suggested six factors, the scree plot showed that two of the factors were close to the line of eigenvalues generated by simulated and resampled random data, and thus, four factors were retained (Figure 2). The four factors were classified as 1) intrinsic reading goals, 2) extrinsic reading goals and learning strategies, 3) desire to understand and reading regulation strategies, and 4) desired reading difficulty.

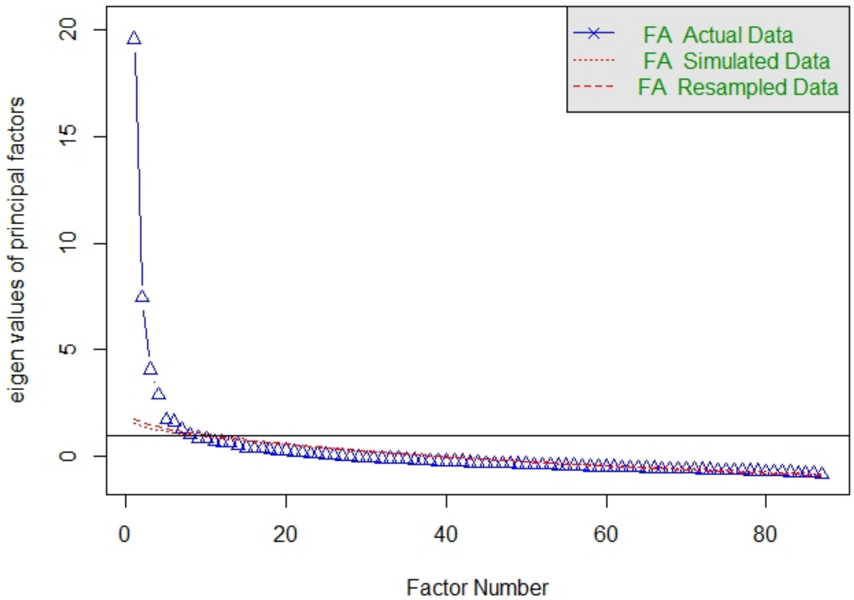

Figure 2. Scree Plot of Parallel Analysis

Next, a PAF estimating four factors using an oblique, oblimin rotation was conducted using the laavan package in R (Rosseel, 2012). Forty-five of the 87 items had factor loadings greater than .5 (Appendix C). After removing all non-significant items, the PAF analysis was conducted again. This process was repeated until all factor loadings were > .5 and cross-loading differences were greater than .15. The final measure contained 31 items (Table 2). The four factors accounted for 51.97% of the variance in the measure. Intrinsic reading goals accounted for 14.48% of variance (7 items), extrinsic reading goals and learning strategies accounted for 12.29% of the variance (8 items), desire to understand and reading regulation strategies accounted for 13.19% of the variance (9 items), and desired reading difficulty accounted for 12.0% of the variance (7 items). This model had a good fit to the data, $\chi^2 = 277.3$, $df = 347$, $p < .999$; $TLI = .902$; $RMSEA = .055$ (90% CI = .05-.06).

Table 2. Reader-based Standards of Cohernece Factor Loadings

Item	Factor 1	Factor 2	Factor 3	Factor 4	Communality
Factor 1: Intrinsic reading goals					
I most often read to relax.	**0.82**	0.15	-0.19	-0.05	0.58
I read because I enjoy it.	**0.81**	-0.05	0.15	0.01	0.79
It is important for me to read for fun, despite my other obligations.	**0.73**	0.08	0.03	0.1	0.62
I feel that it is important to read regularly.	**0.69**	0.03	0.2	0.09	0.69
Reading is not my idea of fun. *	**0.65**	-0.20	0.01	0.28	0.71
One reason I like to read is that it allows me to exercise my imagination and creativity.	**0.63**	0.02	0.16	0.06	0.54
Understanding what I read is part of the fun of reading.	**0.51**	0.01	0.35	0.08	0.57
Factor 2: Extrinsic reading goals and learning strategies					
I routinely make outlines to help me understand what I read.	0.08	**0.81**	-0.13	-0.02	0.68
I often highlight or underline words to follow an author's line of thinking.	0.14	**0.74**	-0.08	0.00	0.57
I often make summaries of what I read.	0.02	**0.71**	-0.1	-0.02	0.53

When I read, I typically take note of key words or concepts in each paragraph.	-0.02	**0.70**	0.11	0.04	0.49
I read so I can have informed conversations with peers/co-workers.	-0.11	**0.60**	0.26	0.10	0.42
I read to study for exams or a job promotion.	-0.17	**0.55**	0.14	-0.03	0.34
I read so I can show others that I'm knowledgeable about a topic.	-0.02	**0.53**	0.19	0.07	0.31
One reason I read is to improve my reasoning skills.	-0.01	**0.52**	0.29	0.23	0.41
Factor 3: Desire understand and reading regulation strategies					
I try to really understand a text, so I can form my own opinions about the content of the text.	0.14	0.07	**0.67**	0.07	0.58
I read to learn more about a topic I'm unfamiliar with.	-0.11	0.03	**0.66**	0.05	0.40
If a text is difficult, I often re-read sections of the text.	0.16	-0.05	**0.63**	-0.19	0.46
When I read, I nearly always want to understand what I've read.	0.27	-0.12	**0.62**	-0.11	0.55
I often look up concepts I do not know.	0.07	0.13	**0.61**	0.13	0.48
When I don't know words in a text, I try to figure out the meaning based on surrounding information.	-0.03	0.01	**0.57**	-0.05	0.31
If I really want to comprehend a book/article, I try to put myself in an environment that will help me do so (e.g., find a quiet place).	0.05	0.09	**0.55**	-0.12	0.31
One benefit of reading is that it can improve my vocabulary.	0.33	0.06	**0.53**	-0.02	0.51
I'm usually more concerned about finishing a text than comprehending it. *	-0.1	-0.34	**0.51**	0.29	0.50
Factor 4: Desired reading difficulty					
I'd rather read easy texts than difficult ones. *	-0.02	0.08	-0.02	**0.80**	0.62
I generally do not like reading dense texts. *	0.08	0.1	-0.13	**0.78**	0.63
I like to read texts that don't require much effort to figure out. *	-0.11	0.09	0.12	**0.75**	0.54
I'd rather read short texts than long ones. *	0.21	0.02	-0.1	**0.62**	0.50
I'd rather read a summary than spend effort understanding a long text. *	0.25	-0.24	0.05	**0.60**	0.63
I try to avoid situations where there is a high chance I will have to read a text carefully. *	0.25	-0.26	0.13	**0.51**	0.59
I often skim when I do not need to know the content very well. *	0.02	-0.07	-0.1	**0.51**	0.26

Note. Bolded items indicate significant factor loadings (> .5). * Denotes reverse-scaled items

The internal consistency for the overall measure was good (α = .89, 95% *CI* = .86 - .90). The sub-scales of intrinsic reading goals (α = .91, 95% *CI* = .89 - .93), extrinsic reading goals and learning strategies (α = .85, 95% *CI* = .82 - .88), desire for understanding (α =.86, 95% *CI* = .82 - .88), and desired reading difficulty (α = .87, 95% *CI* = .84 - .89) also had good internal consistency.

The possible score range of the finalized reader-based standards of coherence measure was -93 to 93, and within the dataset scores ranged from -41 to 85 (*M* = 29.72, *SD* = 22.90). Except for the extrinsic reading goals and learning strategies sub-scale, the sub-scales were lowly to moderately correlated (Table 3).

Table 3. Reader-based Standards of Coherence Factor Correlations

	Factor 1 (7 items)	**Factor 2 (8 items)**	**Factor 3 (9 items)**	**Factor 4 (7 items)**
Factor 1: Intrinsic reading goals	—			
Factor 2: Extrinsic reading goals and learning strategies	-0.01	—		
Factor 3: Desire to understand and reading regulation strategies	0.36*	-0.04	—	
Factor 4: Desired reading difficulty	0.43*	-0.03	0.22*	—

Note. * $p < .05$

2.2.2 Reader-based Standards of Coherence Predicting Reading Habits

Hierarchical regressions tested whether reader-based standards of coherence accounted for variance in people's reading habits when accounting for need for cognition (NFC). A total reader-based standards of coherence score (RB-SOC) was calculated as the sum of scores on the final 31 items. In all regression analyses, Age, Gender, Ethnicity (White vs. Non-white), and Education served as control variables. In separate analyses, both NFC, β = .07, $t(195) = 6.18$, $p < .001$; $R^2 = .22$, $p < .001$ and RB-SOC, β = .10, $t(195) = 10.04$, $p < .001$; $R^2 = .39$, $p < .001$) significantly predicted reading habits (Table 4).

Table 4. NFC and RB-SOC Predicting Reading Habits in Separate Regression Models

	β	Std. Error	t-value	*p*	
NFC Model					
(Intercept)	8.79	1.36	6.48	0.000	***
Age	0.03	0.03	1.18	0.238	
NFC	0.07	0.01	6.18	0.000	***
SexMale	-0.95	0.50	-1.91	0.057	
Non-White	0.86	0.62	1.37	0.171	
Associate Degree	0.64	1.11	0.58	0.563	
Bachelor's Degree	1.35	0.89	1.51	0.132	
Master's/Doctoral Degree	1.41	1.03	1.37	0.172	
Some College	0.52	0.92	0.56	0.574	
RB-SOC Model					
(Intercept)	7.68	1.21	6.34	0.000	***
Age	0.02	0.03	0.81	0.421	
RB-SOC	0.10	0.01	10.04	0.000	***
SexMale	-0.26	0.44	-0.60	0.551	
Non-White	0.13	0.56	0.24	0.814	
Associate Degree	-0.59	0.99	-0.59	0.553	
Bachelor's Degree	0.32	0.80	0.39	0.695	
Master's/Doctoral Degree	0.47	0.92	0.52	0.607	
Some College	0.12	0.82	0.14	0.887	

Note. NFC = Need for cognition; RB-SOC = Reader-based standards of coherence, $^{***}p < .001$. White, Female, and High-school diploma (or GED) served as the intercept.

In a model that included both NFC and RB-SOC predicting reading habits, RB-SOC remained a significant predictor of reading habits; however, NFC no longer predicted reading habits (Table 5). The overall model accounted for a significant amount of variance in reading habits ($R^2 = .39$, $p < .001$). When the combined model was compared to NFC alone, the change in R^2 was significant, $F(1, 194) = 52.34$, $p < .001$; $\Delta R^2 = .17$. However, the change in R^2 was not significant when the combined model was compared to RB-SOC alone, $F(1, 194) = .19$, $p = .661$; $\Delta R^2 < .01$. Thus, when controlling for NFC, RB-SOC predicted reading habits; however, NFC did not predict reading habits when accounting for RB-SOC.

Table 5. NFC and RB-SOC Predicting Reading Habits in a Single Regression Model

	β	Std. Error	t-value	*p*	
(Intercept)	7.72	1.22	6.34	0.000	***
NFC	0.01	0.01	0.44	0.661	
RB-SOC	0.10	0.01	7.24	0.000	***
Age	0.02	0.03	0.83	0.407	
Sex Male	-0.30	0.45	-0.67	0.501	
Non-White	0.17	0.56	0.30	0.766	
Associate Degree	-0.55	1.00	-0.55	0.582	
Bachelor's Degree	0.31	0.81	0.38	0.704	
Master's/Doctoral Degree	0.46	0.92	0.50	0.621	
Some College	0.10	0.82	0.13	0.899	

Note. NFC = Need for cognition; RB-SOC = Reader-based standards of coherence, $R^2 = .39$; ***$p < .001$. White, Female, and High-school diploma (or GED) served as the intercept.

A regression analysis tested whether the demographic factors of Age, Gender, Race (White vs. Non-white), and Education level predicted RB-SOC. Of these variables, only Education level significantly predicted RB-SOC (Table 6). Participants with an Associate’s degree, Bachelor’s degree, and Master’s/Doctorate degree scored higher on the RB-SOC measure than participants with a high-school diploma (or GED).

Table 6. Demographic Information Predicting RB-SOC

	β	Std. Error	t-value	*p*	
(Intercept)	12.52	8.62	1.45	0.148	
Age	0.09	0.19	0.47	0.640	
SexMale	-3.70	3.14	-1.18	0.240	
Non-White	5.30	3.95	1.34	0.181	
Associate Degree	14.13	7.03	2.01	0.046	*
Bachelor's Degree	20.89	5.55	3.76	0.000	***
Master's/Doctoral Degree	20.76	6.40	3.24	0.001	**
Some College	10.17	5.83	1.74	0.083	

Note. RB-SOC = Reader-based standards of coherence; White, Female, High-school diploma (or GED) serve as the intercept. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 7 shows the correlations among the four RB-SOC sub-scales and reading habits. Similar to the factor analysis, the extrinsic reading goals and learning strategies sub-scale did not correlate with the other three sub-scales of the RB-SOC measure. All four RB-SOC sub-scales significantly correlated with reading habits (ARHQ). Intrinsic reading goals ($r = .56$) and desired reading difficulty ($r = .45$) had the two strongest correlations with reading habits.

Table 7. Correlations among RB-SOC Sub-scales and Reading Habits

	1	2	3	4	5	6
1. NFC	--					
2. Reading Habits	0.42***	--				
3. Intrinsic reading goals	0.44***	0.56***	--			
4. Extrinsic reading goals	0.20***	0.27***	0.05	--		
5. Desire to understand	0.39***	0.32***	0.51***	0.05	--	
6. Desired reading difficulty	0.69***	0.45***	0.56***	-0.01	0.29***	-

Note. RB-SOC = Reader-based standards of coherence, NFC = Need for cognition; *** $p < .001$, ** $p < .01$

2.3 Discussion

This study created and tested a measure of reader-based standards of coherence. Importantly, the study demonstrated that this measure was a good predictor of reading habits beyond need for cognition, which is a more general measure of engaging in complex cognitive tasks. Initially, I hypothesized that there would be three sub-categories related to reader-based standards of coherence: typical reading goals, desire to understand, and reading comprehension strategies. Rather than these three sub-categories, analyses suggested that four factors would best describe the data: intrinsic reading goals, extrinsic reading goals and learning strategies, desire to understand and reading regulation strategies, and desired reading difficulty.

Similar to other motivation scales (Henk & Melnick, 1995; Kingston, Wang, Davis, Tonks, Tiemann, & Hock, 2017; Schutte & Malouff, 2007; Wigfield & Guthrie, 1997), the reader-based standards of coherence measure distinguished intrinsic from extrinsic reading goals. The extrinsic goals sub-scale included items about learning strategies associated with extrinsic goals of learning for work or school, such as highlighting, underlining, and making outlines. This suggests that strategies readers employ to learn or memorize material for an extrinsic reading goal may be different than those used for intrinsic reading goals. Importantly, in the present study extrinsic reading goals and learning strategies were uncorrelated with the other three sub-scales. This finding suggests that readers' own motivations for reading may not necessarily align with extrinsic reading goals. This finding somewhat aligns with findings from the AMRS (Schutte & Malouff, 2007), which showed that the reading for recognition sub-scale (e.g., "It is important for to me to have others remark on how much I read") had low correlations with reading as part of self, reading efficacy, and reading to do well in other realms (*r*s = .15 – .16). However, in their study, Schutte and Malouff also found that reading to do well in other realms (e.g., "I read to improve my work or university performance”, “I do all the expected reading for work or university courses”) was modestly correlated with reading as part of self (r = .47) and reading efficacy (r = .36). Because reading to do well in other realms closely aligns with the current study’s extrinsic reading goals and learning strategies sub-scale, it is surprising that these extrinsic goals were not at least lowly correlated with the other sub-scales.

One explanation for these differences might be a difference in sample compositions. Schutte and Malouff (2007) included adults from the general community and a university, with university students comprising about two-thirds of the sample. College students—who have monetary and career incentives to perform well in their classes—may internalize extrinsic reading goals more than non-students, resulting in a stronger relationship between extrinsic and intrinsic reading goals. The present study recruited participants online via MTurk, and although current status (student vs. non-student) was not assessed, it is likely that a majority of the participants were not enrolled in school when they completed the study. In the present study, 74% of participants indicated they had a high-school diploma, Associate Degree, Bachelor’s Degree, or Master’s/Doctorate degree, whereas only 26% indicated they had completed some college.

The third sub-scale, desire to understand and reading regulation strategies, taps into the core of standards of coherence, setting a threshold for understanding (van den Broek et al.,1995, 2001). Reading regulation strategies are used when comprehension fails and are a result of comprehension monitoring. Although the items designed to explicitly measure comprehension monitoring (e.g., “When I read, I frequently check to make sure I'm understanding the text.”) did not have significant factor loadings in the PAF analysis, items related to reading regulation strategies that result from identifying a comprehension failure did have significant factor loadings. Readers may be better at explicitly noticing the strategies that result from comprehension monitoring more than noticing the act of comprehension monitoring itself. Indeed, much research on comprehension monitoring assesses it via inconsistency detection indicated by slower reading times (e.g., Albrecht & O'Brien, 1993; Kim et al., 2018; O'Brien et al., 1998) or asking participants to indicate if there was an inconsistency in a passage (Cain, Oakhill, & Bryant, 2004). Overall, employing reading regulation strategies relates to one's desire to understand texts and is most likely a reader-initiated process (van den Broek & Helder, 2017).

Desired reading difficulty was the final sub-scale of the reader-based standards of coherence measure. Individuals who enjoy reading difficult texts and like experiencing reading challenges are likely to have high levels of text understanding. This sub-scale reflects a readers' need to engage with difficult texts and had a high correlation with need for cognition (r = 69).

In summary, the results demonstrate that the reader-based standards of coherence measure is a more fine-grain measure of reading habits than need for cognition. One benefit of this measure is that it incorporates aspects of reading motivation along with a desire to understand texts, which relates most closely to standards of coherence. The reader-based standards of coherence measure also provided a desired reading difficulty sub-scale which may be interpreted as a reading-specific measure of need for cognition. Overall, reader-based standards of coherence predicted readers' reading habits, but how do these standards fit into a model of reading comprehension? Study 2 used structural equation modeling to answer this question and to determine whether reader-based standards of coherence have a direct effect on reading comprehension.

3.0 Study 2: A Model of Reading Comprehension

The primary aim of Study 2 was to position the reader-based standards of coherence measure in a model of reading comprehension. Using a sample of college students, Study 2 also sought to replicate the finding from Study 1 that reader-based standards of coherence significantly predicts readers' reading habits beyond need for cognition. Because Study 1 used a community sample, it is possible that reader-based standards of coherence measured from a university sample in Study 2 will show different interrelationships among the sub-scales (intrinsic reading goals, extrinsic reading goals and learning strategies, desire to understand and reading regulation strategies, and desired reading difficulty).

Another aim of Study 2 was to test whether comprehension differences between more- and less-skilled readers (e.g. Braze et al., 2007) are reading specific, or if they extend to listening comprehension as well. The study also assessed differences between literal and inferential comprehension in both reading and listening modalities. Do less-skilled readers experience difficulties with both literal and inferential comprehension? Perhaps they experience greater difficulties with one type of comprehension skill over the other. The SEM included measures of literal and inferential comprehension to show their overall contribution to comprehension ability.

The model incorporated relationships among reading abilities predicted by the SVR (Gough & Tunmer, 1986, Tunmer & Chapman, 2012) and the Reading Systems Framework (Perfetti & Stafura, 2014). These two frameworks have differing assumptions regarding the role of vocabulary knowledge. Whereas the SVR assumes that vocabulary knowledge is encapsulated within listening comprehension, the Reading Systems Framework positions it as the link between decoding and reading comprehension, with general knowledge influencing vocabulary knowledge and reading comprehension. Using structural equation modeling, a recent adaptation of the SVR has also demonstrated that background knowledge is a significant predictor of reading comprehension (Talwar, Tighe, & Greenberg, 2018). Finally, although the Reading Systems Framework includes bidirectional links between reading processes and knowledge sources, the SEM tested solely unidirectional relationships because reading skills were assessed at a single time point.

3.1 Methods

3.1.1 Participants

The study included 187 participants recruited through the University of Pittsburgh's psychology subject pool. All participants received course credit for their time. Participants had an age range of 18-56 (*M* = 19.09; *SD* = 3.43); 64 were male and 123 were female. Participants completed a series of reading-related assessments over the course of three sessions.

3.1.2 Materials

3.1.2.1 Reader-based Standards of Coherence and Need for Cognition

Participants completed a modified version of the 31-item reader-based standards of coherence measure from Study 1. In the original version, all seven items in the desired reading difficulty sub-scale were reverse scaled and modifications were made such that only four of these items were reverse scaled and the other three sub-scales had at least two reverse scaled items

(Appendix D). Out of the 31 items, 11 were now reverse scaled. Participants took the same need for cognition questionnaire as Study 1.

3.1.2.2 Comprehension Assessments

Participants completed reading and listening comprehension tasks. To allow for parallel measures of listening and reading comprehension, participants completed two versions of the Gates-MacGinitie reading test, one for listening and one for reading comprehension (Level AR; MacGinitie, MacGinitie, Marie, & Dreyer, 2000). For the two forms of the Gates-MacGinitie comprehension test (Forms T and S), audio versions of the tests were recorded. The forms for listening and reading modalities were randomized across participants. One property of listening tasks is that aurally presented information typically is not repeated. To increase the similarity between reading and listening tasks, participants could not refer to the passages when answering the multiple-choice questions during reading. Both versions of the comprehension test comprised of 11 passages with 48 multiple-choice questions. Each question had four options and 20 of the questions tested literal comprehension while the other 28 tested inferential comprehension. In the current study, separate scores were calculated as percent correct for literal and inferential questions. Participants had a maximum of 35 minutes to complete each comprehension test.

For an additional measure of reading comprehension, participants completed the comprehension component of the Nelson-Denny test (Form G; Brown, Fishco, & Hanna, 1993). The reading comprehension test consisted of seven passages and 38 multiple-choice questions, with five answer options each. Readers could refer to the material as they answered questions and had 20 minutes to complete the task. Half of the questions assessed literal comprehension and the other half assessed inferential comprehension. Similar to the Gates-MacGinitie comprehension test, separate scores were calculated as percent correct for literal and inferential questions.

3.1.2.3 Vocabulary Assessments

Participants also completed three measures of vocabulary knowledge: the vocabulary portion of the Gates-MacGinitie test (McGinitie et al., 2000), the Nelson-Denny test (Form G; Brown, Fishco, & Hanna, 1993) and the Peabody Picture Naming Vocabulary Test (PPVT-4; Dunn & Dunn, 1981). The vocabulary section of the Gates-MacGinitie consisted of 45 multiple-choice items with five options each. Participants selected the word or phrase that best described an underlined word given in a prompt within a 20-minute time limit. The maximum score for this test was 45. The Nelson-Denny vocabulary test has 80 multiple-choice items with five options each. Participants read a prompt and selected the word or phrase that described an italicized word within a 15-minute time limit. The maximum score for the test was 80. The PPVT-4 required participants to select a picture, from an array of four, that described a spoken word. The task included 19 sets of 12 items which increased in difficulty as the set number increased. Participants started on sets according to their age, 17-18 (set 13) or 19+ (set 14). Raw scores were calculated by subtracting the ceiling item (highest item in a set with fewer than 8 errors) from the total number of errors. Raw scores were then converted to standardized scores (M = 100) to account for age and ceiling item differences. Using both visual and aural vocabulary measures should capture the richer receptive vocabulary associated with reading while also including a vocabulary measure that does not rely on decoding ability (Braze et al., 2007; National Reading Panel; 2000).

3.1.2.4 Reading Experience

As one measure of reading experience, participants completed the same reading habits questionnaire (ARHQ) administered in Study 1. As another measure of reading experience, participants also completed three versions of an author recognition test (ART). In the ART, participants viewed a list of names, some of which were authors, and selected names they recognized as authors. ART-1989 was developed by Stanovich and West (1989) and contained 50 authors and 30 foils. ART-2002 is a modified version of Stanovich and West's ART (Perfetti & Hart, 2002) and contained 40 authors and 40 foils. ART-2008 was created by Acheson, Wells, and MacDonald (2008) and contained 65 authors and 65 foils. Across the three tests, six authors that appeared in ART-1989 also appeared in ART-2002, four authors that appeared in ART-2002 appeared in ART-2008, and 23 authors that appeared in ART-1989 also appeared in ART-2008[1]. The three versions were administered as a single test; however, scores were calculated separately for each version. Participants were given a list of 261 names, and across the three lists, there were 126 authors and 135 foils. The foils were names of famous singers, actors, politicians, athletes, and researchers. Scores for each version were calculated as correct minus incorrect. Score ranges for each version are as follows: ART-1989 -30 to 50; ART-2002 -40 to 40; ART-2008 -65 to 65. The three different versions allow for an assessment of whether using older or newer authors influences scores on reading experience measures and whether they have different relationships to reading abilities.

3.1.2.5 Decoding

Measures of pseudoword decoding, reading fluency, word-form knowledge, and orthographic knowledge were included as measures of decoding ability. The pseudoword decoding test is a pseudohomophone choice task, in which participants selected which pseudowords sounded like real words (e.g., fite; Olson, Forsberg, Wise, & Rack, 1994). Items included in this test were

---

[1] The three versions of the author recognition test are part of another study examining differences between test versions and their foils, which did not overlap across test versions.

created by Olson, Wise, Conners, Rack, and Fulker (1989) and Perfetti and Hart (2002). The pseudohomophone task was scored using d-prime (z-score transformed hits – z-score transformed false alarms).

The Test of word reading efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) is a measure of reading efficiency in which readers' accuracy for pronouncing real words and pseudowords was assessed. Participants had 45 seconds to read 104 words as fast as they could. They had another 45 seconds to read 63 pseudowords as fast as they could. The test has two scores, one for pseudoword reading and another for word reading. Raw scores were calculated as hits minus incorrect/skipped words. Scores were then standardized using the scoring manual (Torgesen et al., 1999).

A word-form manipulation test required participants to insert and delete specified phonemes and served as an assessment of readers' word-form knowledge. In this test, 19 items were provided, and participants added or removed a specified sound of a given word and wrote the new word (Olson et al., 1989; Perfetti & Hart, 2002). For example, if participants were instructed to remove the /p/ sound from "speak", the resulting word would be "seek". Next, if participants were instructed to add the /l/ sound in the position where they removed the first sound, the new word would be "sleek". Adding full (correct spelling, phonology, and word status) and partial (only correct phonology, spelling, or word status) points provides the total score on this task (reported as a percentage out of 114). This task relies on both phonological and orthographic knowledge.

In a test of orthographic knowledge, participants viewed a list of 140 letter-strings and indicated which words from the list had correct spellings. Items in this test came from the Baroff spelling test (Olson et al., 1989; Perfetti & Hart, 2002). The task was scored using d-prime.

3.2 Procedure

Participants completed the study over the course of three sessions. In Session 1, participants first completed the reader-based standards of coherence, need for cognition, and reading habits measures. After an intervening task used in Study 3, participants completed listening/reading versions of the Gates-MacGinitie comprehension test. This was counterbalanced, such that half of the participants received the reading version first. Next, participants completed the Gates-MacGinitie vocabulary test, followed by the author recognition test and spelling test. Participants then completed the listening/reading version of the Gates-MacGinitie comprehension test in the modality that they had not completed earlier. The session ended with the pseudoword decoding task.

Participants returned for Session 2 one to four days after Session 1. In Session 2, participants first completed the Nelson-Denny vocabulary test, followed by the Nelson-Denny comprehension test. They then completed the word-form manipulation test, TOWRE word and pseudoword reading tests, and the PPVT-4.

Session 3 was completed 13 to 14 days after Session 2 (14 – 21 days after Session 1). The time delay allowed ample time to assess test-retest reliability of the reader-based standards of coherence measure. The RB-SOC measure was administered in the same order as Session 1, except the scales were reversed. In Session 1 the scale ranged from strongly disagree to strongly agree and in Session 3 the scale ranged from strongly agree to strongly disagree. Participants then completed two additional tasks used in Study 3.

3.2.1 Analysis Procedure

The first analysis for Study 2 focused on replicating findings from Study 1 that the reader-based standards of coherence measure was a predictor of readers' reading habits beyond need for cognition. Test-retest reliability of the measure over a 14- to 21-day time span was also assessed.

The second analysis focused on the goal of positioning reader-based standards of coherence within a structural equation model of reading comprehension. First, a measurement model was estimated with six latent factors: Decoding, Vocabulary Knowledge, Listening Comprehension, Reading Comprehension, Reading Experience, and Reader-based Standards of Coherence (first administration). Indicators with non-significant factor loadings were removed and not entered in the full structural equation model.

The SEM incorporates assumptions made by the Simple view of reading (Gough & Tunmer, 1986; Tunmer & Chapman, 2012) and Reading Systems Framework (Perfetti, 1999; Perfetti & Stafura, 2014). Rather than assuming that one framework might best explain the covariance structure of the reading data, I tested the assumptions of both the SVR and Reading Systems Framework within a single model. The SVR predicts that both decoding and listening comprehension will predict reading comprehension. The SVR also predicts that vocabulary knowledge should predict listening comprehension ability. The Reading Systems Framework predicts that rather than decoding directly predicting reading comprehension, decoding predicts vocabulary knowledge, which then predicts reading comprehension ability. Past research has also found that reading experience (Landi, 2010; Stanovich & Cunningham, 1992; Stanovich & West, 1997) and background knowledge (e.g., Talwar et al., 2018) predict reading comprehension, therefore this path was included in the model. Finally, the model included paths from reader-based standards of coherence to reading experience and reading comprehension. Non-significant regressions between latent factors were removed, one at a time, to refine the model.

3.3 Results

Of the 187 participants, 178 completed Sessions 1 and 2. Because all measures were required for the SEM analysis and data were not missing at random, only participants who completed both sessions were retained. Scores on all measures can be found in Table 8. Unlike Study 1, the extrinsic reading goals and learning strategies sub-scale of reader-based standards of coherence was significantly correlated with all other sub-scales (Table 9).

Table 8. Descriptive Statistics of Measures

	M	*SD*
Word-form Manipulation	0.82	0.12
RB-SOC	18.19	21.98
NFC	13.23	15.89
Reading Habits	9.40	3.32
Pseudoword Decoding	2.34	0.65
Spelling	1.90	0.77
PPVT-4	106.43	8.89
ART-1989	4.17	2.46
ART-2002	8.46	4.38
ART-2008	11.97	5.88
Nelson-Denny Vocabulary	66.46	6.46
Nelson-Denny Comprehension	0.86	0.11
TOWRE-Words	95.61	10.37
TOWRE-Pseudowords	100.92	10.98
Gates-MacGinitie Vocabulary	37.97	4.12
Gates-MacGinitie Reading	0.78	0.10
Gates-MacGinitie Listening	0.73	0.12

Note. ART = Author recognition test; PPVT-4 = Peabody Picture Vocabulary Test; TOWRE = Test of word reading efficiency

Table 9. Correlations among RB-SOC Sub-scales

	1	2	3	4
1. Intrinsic reading goals	---			
2. Extrinsic reading goals and learning strategies	0.25*	---		
3. Desire to understand and reading regulation strategies	0.51*	0.40*	---	
4. Desired reading difficulty	0.64*	0.31*	0.45*	---

Note. * $p < .05$

3.3.1 Study 1 Replication and Test-retest Reliability

To replicate previous findings (Study 1) that reader-based standards of coherence (RB-SOC) predicted reading habits beyond need for cognition (NFC), I conducted linear regressions predicting reading habits from need for cognition and reader-based standards of coherence with Age and Gender as control variables. The internal consistency for RB-SOC was good, α = .88. In separate analyses, both NFC, $R^2 = .10$, $p < .001$ and RB-SOC, $R^2 = .44$, $p < .001$ significantly predicted readers' reading habits (Table 10). When both NFC and RB-SOC were included in a single model, NFC no longer predicted reading habits, $R^2 = .49$, $p < .001$. This replicates findings from Study 1 that the RB-SOC measure is a better predictor of readers' reading habits than NFC. Furthermore, RB-SOC had good test-retest reliability ($r = .89$) with a time interval of 14 – 21 days between the first and second administration of the measure. The internal consistency for the second administration was also good, α = .88.

Table 10. RB-SOC and NFC Predicting Reading Habits: Study 2

	β	Std. Error	t-value	*p*	
NFC Model					
(Intercept)	6.50	1.36	4.79	0.000	***
Age	0.11	0.07	1.59	0.113	

SexMale	0.13	0.52	-0.25	0.806	
NFC	0.06	0.02	4.19	0.000	***
RB-SOC Model					
(Intercept)	5.29	1.07	4.95	0.000	***
Age	0.11	0.05	1.99	0.048	*
SexMale	0.71	0.40	1.78	0.078	
RB-SOC	0.10	0.01	11.62	0.000	***
Combined NFC and RB-SOC Model					
(Intercept)	5.36	1.08	4.98	0.000	
Age	0.10	0.05	1.95	0.053	
SexMale	0.79	0.42	1.89	0.063	
RB-SOC	0.10	0.01	8.97	0.000	***
NFC	-0.01	0.01	-0.68	0.500	

Note. NFC = Need for cognition; RB-SOC = Reader-based standards of coherence, *p < .05, ***p < .001

3.3.2 Skill Differences in Reading and Listening Comprehension

To test whether reading skill influenced scores on both reading and listening comprehension measures (e.g., Braze et al., 2007), a linear regression tested the effects of Reading Skill, Question Type, and Modality on comprehension (Table 11). The Gates-MacGinitie test was used as a continuous measure to define reading skill.

Table 11. Linear Regression on Reading Skill Differences: Gates-MacGinitie

	β	Std. Error	t value	p	
(Intercept)	0.10	0.03	3.80	0.000	***
Reading skill (Gates-MacGinitie)	0.82	0.03	25.22	0.000	***
Modality	-0.12	0.03	-4.74	0.000	***
Question Type	0.27	0.05	5.30	0.000	***
Reading Skill x Modality	0.20	0.03	6.24	0.000	***
Reading Skill x Question Type	-0.19	0.06	-2.90	0.004	**
Modality x Question Type	0.01	0.05	0.29	0.774	
Reading skill x Modality x Question type	-0.04	0.06	-0.63	0.532	

Note. **p < .01, ***p < .001.

As reading skill increased, participants performed better on both types of comprehension questions (Table 11). Participants also performed better on questions in the reading than the listening modality and they performed better on inferential questions than literal questions (Table 12). The Reading Skill x Modality and Reading Skill x Question Type interactions were significant (*p*s < .005), and the Modality x Question Type and Reading Skill x Modality x Question Type interactions were not significant (*p*s > .532). As reading skill increased, participants performed better in the reading modality than the listening modality (Figure 3). Additionally, as reading skill increased, scores on the literal questions became more similar to scores on the inferential questions (Figure 4).

Table 12. Comprehension Differences by Question Type

	Inference	Literal
	M (*SD*)	*M* (*SD*)
Listening: Gates-MacGinitie	71.83 (13.60)	62.75 (14.58)
Reading: Gates-MacGinitie	82.50 (10.60)	76.83 (11.29)
Reading: Nelson-Denny	78.71 (14.23)	93.49 (9.40)

Note. Values represent percent correct on comprehension tests.

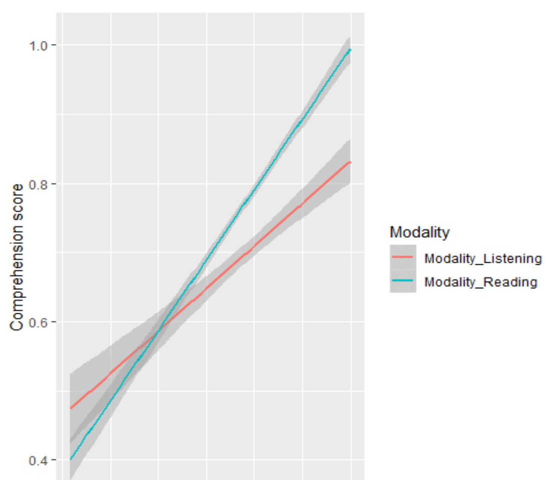

Figure 3. Reading Skill x Modality Interaction: Gates-MacGinitie

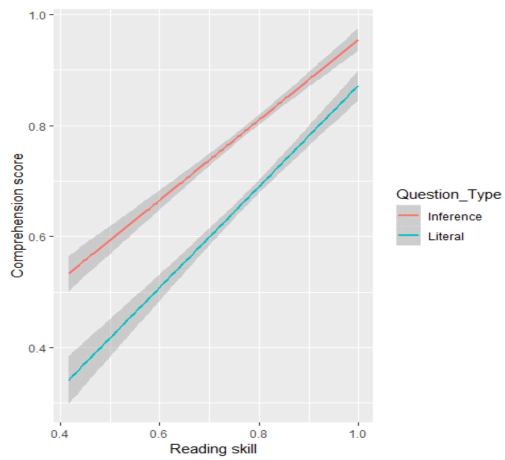

Figure 4. Reading Skill x Question Type Interaction: Gates-MacGinitie

Interestingly, differences in accuracy between the literal and inferential questions on the Nelson-Denny comprehension test showed a different pattern. An additional linear regression analysis tested effects of Reading Skill (assessed by the Gates-MacGinitie test) and Question Type on comprehension scores on the Nelson-Denny test. For the Nelson-Denny assessment readers could refer to the passages while answering multiple-choice questions.

Table 13. Linear Regression on Reading Skill Differences: Nelson-Denny

	β	Std. Error	t value	p	
(Intercept)	0.45	0.04	10.67	0.000	***
Reading skill (Gates-MacGinitie)	0.53	0.05	9.87	0.000	***
Question Type	0.23	0.04	5.57	0.000	***
Reading Skill x Question Type	-0.21	0.05	-3.85	0.000	***

Note. *** $p < .001$.

As Reading Skill increased, participants performed better on both literal and inferential questions (Table 12). Additionally, participants responded more accurately to literal questions than inferential questions. This pattern is a reversal from performance on the Gates-MacGinitie test in which inferential questions were answered more accurately than literal questions. The Reading Skill x Question Type interaction was also significant; as reading skill increased, the difference between inference and literal questions decreased (Figure 5).

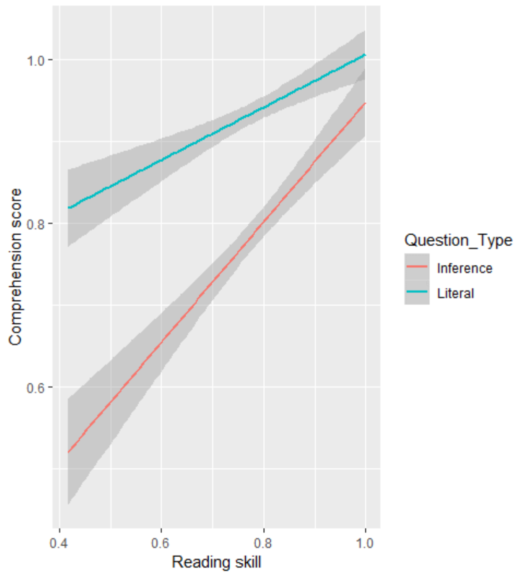

Figure 5. Reading Skill x Question Type Interaction: Nelson-Denny

Thus, whereas readers performed better on inferential questions on the Gates-MacGinitie reading test, they performed better on literal questions on the Nelson-Denny reading test. The discrepancy between the performance on the two tests for literal and inferential questions could be due to task differences; readers were unable to refer to texts when answering questions for the Gates-MacGnitie test but were able to for the Nelson-Denny test. Alternatively, differences in question design may have influenced performance on the two tests. Participants scored better on inferential questions on the Gates-MacGinitie test than inferential questions on the Nelson-Denny test, $t(177) = 3.96$, $p < .001$ (Table 12). Conversely, for literal questions, participants performed better on the Nelson-Denny test than the Gates-MacGinitie test, $t(177) = -20.77$, $p < .001$. Implications for this finding are discussed further in the General Discussion.

3.3.3 Measurement Model

Prior to conducting a structural equation model, a measurement model was estimated with the indicator variables hypothesized to load onto their respective latent variables. Please see Table 13 for the latent factors and their indicator variables. The measurement model consisted of six latent factors of Decoding, Vocabulary Knowledge, Reading and Listening Comprehension, Reading Experience, and RB-SOC. For the Decoding latent factor, the indicators were Word-form Manipulation, Spelling, TOWRE Words, TOWRE Pseudowords, and Pseudoword Decoding. The residual errors for TOWRE Words and Pseudowords were allowed to covary. The Vocabulary Knowledge latent factor included the Nelson-Denny vocabulary test, Gates-MacGinitie vocabulary test, and the PPVT-4. The Reading Comprehension latent factor included inferential and literal questions from the Gates-MacGinitie and Nelson-Denny reading tests and the Listening Comprehension latent factor had indicators from literal and inferential questions on the listening version of the Gates-MacGinitie test. The Reading Experience latent factor included the three versions of the ART (1989, 2002, and 2008) and the reading habits questionnaire. Error terms for the three versions of the ART were allowed to covary because the measures contained overlapping items (hits only). Finally, the RB-SOC latent factor included indicators for intrinsic reading goals, extrinsic reading goals and learning strategies, desire to understand and reading regulation strategies, and desired reading difficulty.

Because some measures had non-normal distributions (Appendix E), the Satorra-Bentler correction was implemented. The Satorra-Bentler correction provides χ^2 values adjusted for kurtosis (Satorra & Bentler, 1994). The χ^2 was significant for all analyses, potentially due to the large number of parameters estimated, and the small sample to parameters ratio (3:1; Kline, 2011).

Table 14. Measurement Model Factor Loadings

	Standardized factor loading	z-value	p
Listening Comprehension			
Gates-MacGinitie Inferential (Listening)	0.84		
Gates-MacGinitie Literal (Listening)	0.79	11.04	0.000
Reading Comprehension			
Gates-MacGinitie Inferential (Reading)	0.72		
Gates-MacGinitie Literal (Reading)	0.63	7.11	0.000
Nelson-Denny Literal (Reading)	0.60	4.36	0.000
Nelson-Denny Inferential (Reading)	0.76	6.70	0.000
Vocabulary			
Nelson-Denny Vocabulary	0.86		
PPVT-4	0.79	12.30	0.000
Gates-MacGinitie Vocabulary	0.47	4.73	0.000
Decoding			
Pseudoword Decoding	0.66		
Spelling	0.34	3.68	0.000
Word-form Manipulation	0.71	5.53	0.000
TOWRE Words	0.11	1.14	0.255
TOWRE Pseudowords	0.44	4.67	0.000
Reader-based Standards of Coherence			
Intrinsic Reading Goals	0.86		
Extrinsic Reading Goals/Learning Strategies	0.38	4.55	0.000
Desired Difficulty	0.74	10.68	0.000
Desire to Understand/Regulation Strategies	0.60	8.52	0.000
Reading Experience			
Reading Habits	0.70		
ART-1989	0.52	6.54	0.000
ART-2002	0.58	6.89	0.000
ART-2008	0.65	7.40	0.000

Note. Items entered first serve as the scaling variable. PPVT-4 = Peabody Picture Vocabulary Test; ART = Author recognition test; TOWRE = Test of word reading efficiency.

Although the χ^2 was significant, other fit indices indicated that the measurement model had acceptable fit, $\chi^2(190)= 309.08$, $p < .001$, *CFI* = .92, *TLI* = .91, *RMSEA* = .059 (90% CI: .047 = .071). All indicators except TOWRE Words ($p = .255$) significantly loaded onto their latent factors (ps < .001; Table 13). All latent factors were correlated with one another (rs = .33 - .90, ps < .01) except for RB-SOC and Decoding (p= .063). Because TOWRE Words did not have a significant factor loading, the indicator was removed from the model, yielding acceptable model fit, $\chi^2(171)=$ 281.44, $p < .001$, *CFI* = .92, *TLI* = .91, *RMSEA* = .060 (90% CI: .048 = .072).

3.3.4 Structural Equation Model

In the SEM, theoretical assumptions from the SVR and Reading Systems Framework were applied along with predictions for how reader-based standards of coherence would fit into a model of reading comprehension. In accordance with the SVR, Listening Comprehension predicted Reading Comprehension and Vocabulary Knowledge predicted Listening Comprehension. To test the Reading Systems Framework assumptions, Vocabulary Knowledge and Reading Experience predicted Reading Comprehension and Decoding predicted Vocabulary Knowledge. Finally, RB-SOC predicted both Reading Experience and Reading Comprehension (See Figure 6 for the full SEM).

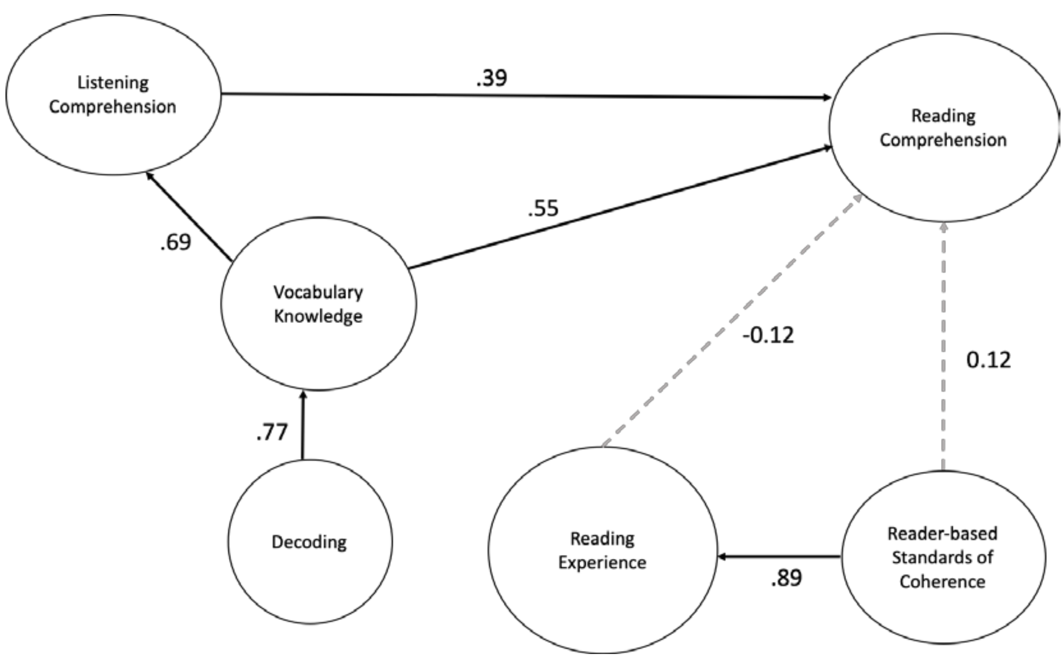

Figure 6. Full Structural Equation Model with all Hypothesized Predictions

The SEM had reasonable fit (Table 15) but paths from Reading Experience and RB-SOC did not have significant regression coefficients. Non-significant paths were removed one at a time until only significant regressions remained. After removing the path from RB-SOC to Reading Comprehension, Reading Experience still did not predict Reading Comprehension (β < .01; *p* > .999; Table 15).

Table 15. SEM Fit Statistics

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI
1. Full SEM model	320.30*	178	0.91	0.89	0.067	0.055 - 0.078
2. RB-SOC - Reading Comprehension removed	320.97*	179	0.91	0.89	0.068	0.056 - 0.080
3. Reading Experience - Reading comprehension removed[a]	296.81*	179	0.92	0.91	0.062	0.049 - 0.074
4. Vocabulary Knowledge – Reading Comprehension removed	316.42*	180	0.91	0.90	0.066	0.054 - 0.078

Note. RB-SOC = Reader-based standards of coherence. a. RB-SOC predicted vocabulary knowledge. * $p < .05$.

According to the Reading Systems Framework, background knowledge should predict both reading comprehension and vocabulary knowledge (Perfetti & Stafura, 2014). A study testing effects of reading experience found that reading experience predicted vocabulary knowledge rather than reading comprehension (e.g., Cunningham & Stanovich, 1997). With Reading experience as one example of background knowledge, a new model included a path from Reading Experience to Vocabulary Knowledge and excluded Reading Experience predicting Reading Comprehension (Figure 7). This model yielded acceptable fit and Reading Experience significantly predicted Vocabulary Knowledge (β = .42, *p* < .001; Table 15). Thus, RB-SOC predicted Reading Experience, which predicted Vocabulary Knowledge. In a mediation analysis, the indirect effect of RB-SOC on Vocabulary Knowledge was significant (indirect effect = .36, *p* < .001). Additionally, Decoding had an indirect effect on Reading Comprehension through Vocabulary

Knowledge (indirect effect = .32, p = .002), but the direct effect was not significant (direct effect = .02, p = .830).

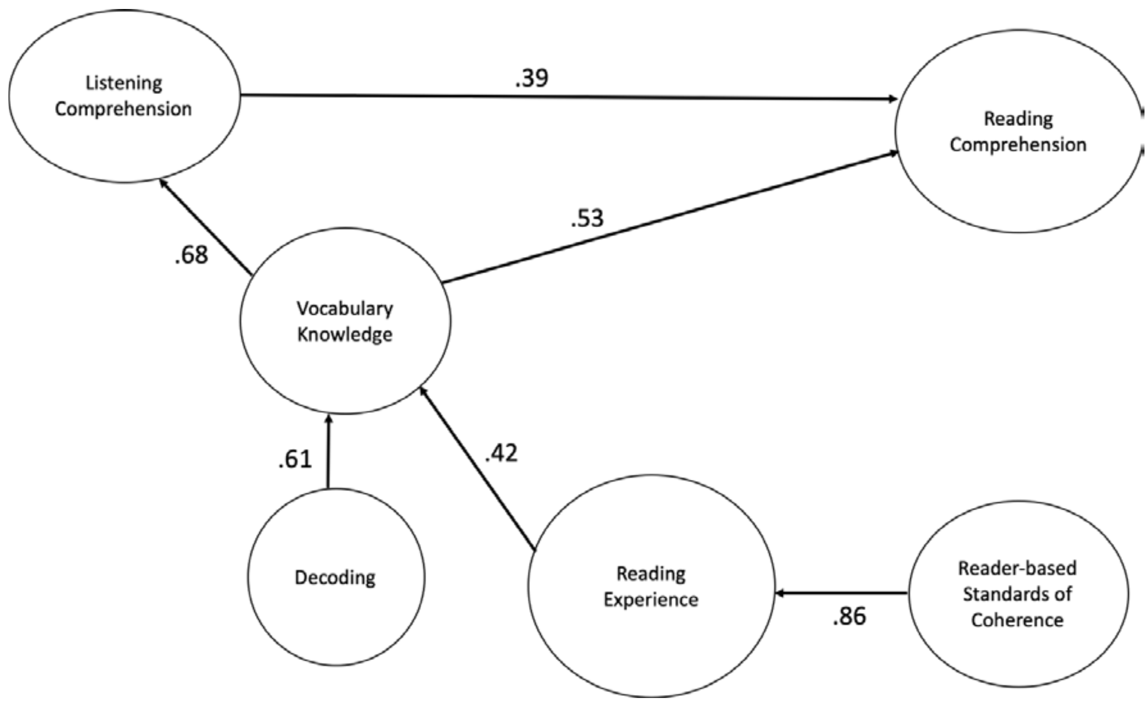

Figure 7. Final Structural Equation Model

Finally, in a direct test of the SVR and the Reading Systems Framework, I tested whether the relationship between Vocabulary Knowledge predicting Reading Comprehension yielded better model fit than a model without Vocabulary Knowledge predicting Reading Comprehension using the χ^2 difference test. The χ^2 difference test assesses changes in χ^2, with significant changes indicating a decrease in model fit. Overall, there was worse model fit when the path of Vocabulary Knowledge predicting Reading Comprehension was removed, $\Delta\chi^2 = 19.61$, $\Delta df = 1$; $p < .001$, $\chi^2 = 316.42$, $df = 180$, $p < .001$, $CFI = .91$, $TLI = .90$, RMSEA = .066 (90% CI = .054 - .078). Thus, vocabulary knowledge accounts for significant variance in reading comprehension when accounting for listening comprehension ability, supporting the Reading Systems Framework, but not the SVR.

3.4 Discussion

Overall, findings demonstrate that the Reader-based standards of coherence measure predicts college students' reading habits, replicating findings from Study 1 with a different sample. The measure also had good test-retest reliability, suggesting that reader-based standards of coherence are relatively stable. Additionally, the SEM demonstrated that reader-based standards of coherence fit into a model of reading comprehension that meets assumptions of both the SVR and Reading Systems Framework. Finally, based on a median split, higher-skilled readers outperformed lower-skilled readers on reading and listening tasks on both literal and inferential questions. First, I discuss the findings of reader-based standards of coherence.

Similar to Study 1, reader-based standards of coherence significantly predicted college students' reading habits beyond need for cognition. Unlike Study 1, all sub-scales of reader-based standards of coherence (intrinsic reading goals, extrinsic reading goals and learning strategies, desire to understand and reading regulation strategies, and desired reading difficulty) were significantly correlated, including extrinsic reading goals and learning strategies, which was uncorrelated to any of the other sub-scales in Study 1. This finding with college students is similar to that of Schutte and Malouff (2007) who found that reading to do well in other realms was modestly correlated with reading as part of self and reading efficacy. As alluded to in the Discussion of Study 1, the relationship between extrinsic reading goals and learning strategies and other sub-scales related to intrinsic motivations for reading may depend on student status (i.e., student vs. non-student). All participants in Study 2 were university students enrolled in an Introduction to Psychology undergraduate course. It is likely that these students may have internalized their extrinsic reading goals more than non-students in Study 1.

In a SEM, all four sub-scales significantly loaded onto a reader-based standards of coherence latent factor. Rather than directly predicting reading comprehension ability, reader-based standards of coherence related to more reading experience, which related to increased vocabulary knowledge. Vocabulary knowledge then predicted readers' reading comprehension ability. This finding somewhat aligns with previous findings of the relationship between reading experience and reading comprehension. Without controlling for vocabulary knowledge, Acheson et al. (2008) found that print exposure (ART-2008 in the present study and a magazine recognition test) was positively related to reading and English ACT scores. Acheson et al. also included a measure that assessed college students' reading habits relative to their peers. The comparative measure predicted performance on reading and English sections of the ACT, but the other measures of print exposure did not. The findings highlight that reading experience, especially when assessed as a comparative measure, is related to reading comprehension.

In a longitudinal study, Cunningham and Stanovich (1997) examined influences on reading ability from first to eleventh grade. When controlling for first grade reading comprehension, vocabulary knowledge, and intelligence scores based on Raven's Matrices, eleventh grade print exposure predicted eleventh grade vocabulary knowledge in addition to cultural and multi-cultural knowledge. However, print exposure did not predict reading comprehension. Other studies have shown similar patterns; reading experience helps strengthen lexical representations at both the word-form and semantic level (Braze et al., 2007; Jenkins, Stein, & Wysock, 1984; Nation, 2017; Nation & Snowling, 1998; Perfetti & Hart, 2001, 2002).

Thus, reading experience contributes to more knowledge of word meanings necessary for reading comprehension. When vocabulary knowledge is not included in models of reading experience predicting comprehension (e.g., Acheson et al., 2008), reading experience may be a significant predictor of reading comprehension. However, this relationship may be because experience affects vocabulary knowledge that facilitates reading comprehension processes (e.g., Perfetti & Hart, 2001). Experiencing a word across multiple contexts helps strengthen its meaning representation (Bolger et al., 2008). The results from the SEM suggest that reader-based standards of coherence influence the reading experiences that readers seek, which in turn aid forming and accessing semantic representations. This finding adds to previous findings that motivation factors are related to reading experiences (Davis et al., 2018; Parault & Williams, 2009). Wigfield and Guthrie (1997) found that children with higher motivation for reading read more frequently and read more books than those with lower reading motivation.

In the present study, listening comprehension ability was a low but significant predictor of reading comprehension when controlling for vocabulary knowledge. This result supports previous SEM findings (Braze et al., 2007; Protopapas, Simos, Sideridis, & Mouzaki, 2012; Talwar et al., 2018; Tunmer & Cunningham, 2012) and the SVR (Gough & Tunmer, 1986). People must have good literal and inferential listening comprehension skills to apply them to reading (e.g., Gough & Tunmer). Many previous studies have used spoken vocabulary assessments to test the SVR. The SVR posits that oral comprehension and decoding ability are two separate components that contribute to reading comprehension (Gough & Tunmer, 1986). The decoupling of reading-specific skills, i.e., decoding from general language ability, is why oral vocabulary measures have been consistently used in tests of the SVR (e.g., Braze et al., 2007, 2016; Protopapas, Simos, Sideridis, & Mouzaki, 2012; Tunmer & Chapman, 2012). However, because reading involves activating word meanings via their visual word-forms, assessing written vocabulary knowledge should also provide unique contributions to reading comprehension, as seen in the present study.

As a test of the SVR, Study 2 also assessed relationships among decoding, vocabulary knowledge, and reading comprehension. Findings from the SEM suggest that for adult college readers, rather than having a direct relationship to reading comprehension, good decoding ability is associated with accessing word meanings, i.e., vocabulary knowledge, which are necessary for successful reading comprehension. This finding supports predictions made by the Reading Systems Framework. Decoding may have a stronger relationship to reading comprehension for younger readers than older readers (e.g., Cromley & Azevedo, 2007; Garcia & Cain, 2014; Kim, 2017) because effort spent on decoding words may leave fewer resources for reading comprehension (e.g., Bell & Perfetti, 1994; Stanovich, 1986). For example, in a SEM with English speaking second graders, Kim (2017) tested a direct and indirect effect model of reading (DIER; Cromley & Azevedo, 2007). In support of the SVR, findings showed that listening comprehension and word reading directly predicted reading comprehension. Importantly, within the DIER model, vocabulary, working memory, comprehension monitoring, inference ability, and theory of mind had indirect effects on reading comprehension through listening comprehension for second graders.

A study with second, third, and fourth grade Greek children demonstrated that the relationship between decoding ability and reading comprehension changed across grade levels (Protopapas et al., 2007). Using measures of decoding fluency, decoding accuracy, oral receptive and expressive vocabulary knowledge, and reading comprehension Protopapas et al. found that across grade levels decoding accuracy and fluency predicted vocabulary knowledge, and vocabulary knowledge predicted reading comprehension. The direct relationship between decoding accuracy and reading comprehension depended on grade level; second graders' decoding accuracy was a stronger predictor of their reading comprehension than fourth graders' decoding accuracy.

The findings from these two studies support the hypothesis that over time readers' decoding ability becomes less directly associated with reading comprehension ability but remains a predictor of vocabulary knowledge. For older readers, especially those who are at a college level, issues with decoding may hinder activating word meanings necessary for comprehension. Thus, the model positions vocabulary knowledge as a focal point in reading comprehension and the indirect relationship between decoding ability and reading comprehension via vocabulary knowledge supports assumptions made by the Reading Systems Framework. Vocabulary knowledge is important for both listening and reading comprehension, and both reading experience and decoding ability directly affect semantic access. Finally, reader-based standards of coherence are one driving factor for readers seeking out these different reading experiences that help them acquire more word meanings.

4.0 Study 3: On-line Inference Generation

Whereas Study 2 assessed how reader-based standards of coherence fit into a model of reading comprehension, Study 3 assessed how reader- and situation-based standards of coherence relate to bridging and predictive inference generation during on-line comprehension. Study 3 predicts that participants with higher reader-based standards of coherence should generate more inferences and read texts more closely, resulting in longer reading times. Situation-based standards of coherence were manipulated by providing participants with one of two comprehension goals: answering open-ended questions vs. phrase matching. Answering open-ended questions should encourage deeper levels of understanding compared to phrase matching, which relies on remembering surface-level text information. Reader- and situation-based standards of coherence may interact; readers with higher reader-based standards of coherence may set higher standards regardless of a particular reading goal, whereas those with lower reader-based standards of coherence may not. Finally, people with higher standards of coherence may express more interest in texts they read/hear during the experiment.

Before determining whether reader-based standards of coherence predict on-line reading comprehension given a particular reading goal, two norming experiments were conducted to refine comprehension materials. The first norming experiment verified that the materials led to hypothesized bridging and predictive inferences. The second norming experiment verified that these materials, which spanned a broad range of topics, generated different levels of interest. This ensures that later interest ratings used during the on-line comprehension tasks would have enough variability to serve as predictors and dependent variables in analyses.

4.1 Methods

4.1.1 Inference Norming and Material Selection

Materials consisted of 19 newspaper articles, ranging in topic (sports, poverty, epidemics, embezzlement, history, and thrillers). These newspaper articles were from the late 19th century to the early 20th century and were selected from *The Washington Times*, *The New York Herald*, and *The Washington Herald*. Newspaper articles were chosen as text materials because they are authentic texts across a variety of topics, offer an opportunity for readers to learn new information, and readers tend to create situation models when they read texts as news stories (Zwaan, 1994). These texts also allow for multiple inference-based sentences without disrupting text-flow. The earlier time period of the articles greatly reduces the likelihood that readers would be familiar with the content and the information should be new to most readers. The newspaper articles had an average length of 681.37 words (SD = 74.42), 38.63 sentences (SD = 4.43), and average Flesch-Kincaid reading grade level of 9.18 (SD = 1.58). For a full list of the newspaper articles and their characteristics, see Appendix F.

Twenty-five university students (M_{age} = 19.63, SD = 1.07; 15 Female, 10 Male) rated how probable it was that an event happened (i.e., bridging inference) or will happen (i.e., predictive inference) in the article using a five-point Likert scale ranging from very unlikely to very likely.

For each newspaper article, participants rated the likelihood of two bridging inferences, two predictive inferences, two events that were very unlikely (i.e., the opposite of likely events), and two events with a moderate likelihood of occurring (i.e., subjective). Participants read the articles in a self-paced, sentence-by-sentence, reading task and occasionally they provided their likelihood rating. All target inferences occurred at the end of sentences. Participants also completed the Gates-MacGinitie comprehension and vocabulary tests (Level AR Form S; MacGinitie et al., 2000) and could refer to the text material when answering the multiple-choice questions in the comprehension test.

4.1.1.1 Inference Norming Results

For the Gates-MacGinitie test, the average comprehension score was 87.42% (*SD* = 6.14%). Readers performed better on inference questions (*M* = 90.03%, *SD* = 8.02%) than literal questions (*M* = 83.54%, *SD* = 9.83%), $t(24) = 2.21$, $p = .037$. The average vocabulary score was 40.60 (*SD* = 2.78).

Each participant's inference ratings were Z-score transformed such that zero became the mean rating for each participant. Inference ratings greater than zero indicated a positive likelihood of occurring whereas ratings lower than zero indicated a negative likelihood of occurring. Linear mixed effects regressions were first conducted using all articles with condition (bridging, predictive, moderate filler, and low filler) predicting inference ratings and subsequent analyses excluded articles in which bridging and/or predictive inferences had ratings lower than zero. Reading grade level and sentence length served as control variables and article nested within subjects served as the random slope in analyses. Treatment coding was employed such that the bridging condition served as the control (intercept) of the models. The bridging condition was chosen as the control condition because bridging inferences are necessary for comprehension (e.g., Grasser et al., 1994) and should have the highest likelihood rating. Overall, all predictors were significant (Table 16). Participants rated bridging inferences as most likely to occur, followed by predictive inferences, moderate fillers (somewhat likely events), and low fillers (unlikely events).

Table 16. Mixed Effects Linear Regressions for Inference Ratings

		β	*df*	*t value*	
Full set of passages n = 19					
	(Intercept)	0.66	3491	8.38	***
	Filler low	-1.56	3657	-43.41	***
	Filler moderate	-0.75	3653	-20.31	***
	Predictive	-0.13	3653	-3.57	***
	Reading grade level	0.02	3650	2.81	**
	Sentence length	-0.02	3675	-7.86	***
Subset of passages n = 16					
	(Intercept)	0.77	3070	9.49	***
	Filler low	-1.55	3055	-40.35	***
	Filler moderate	-0.80	3055	-20.18	***
	Predictive	-0.08	3053	-2.24	*
	Reading grade level	0.02	3063	1.87	
	Sentence length	-0.02	3060	-7.84	***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Reading grade level is calculated as Flesh-Kincaid grade level.

After inspecting the ratings for each condition within each article, three articles had ratings in which the predictive and/or bridging inferences received ratings below zero (Z-scores < -.19). After excluding these three articles, an identical analysis revealed a similar pattern of results (Figure 8; Table 16).

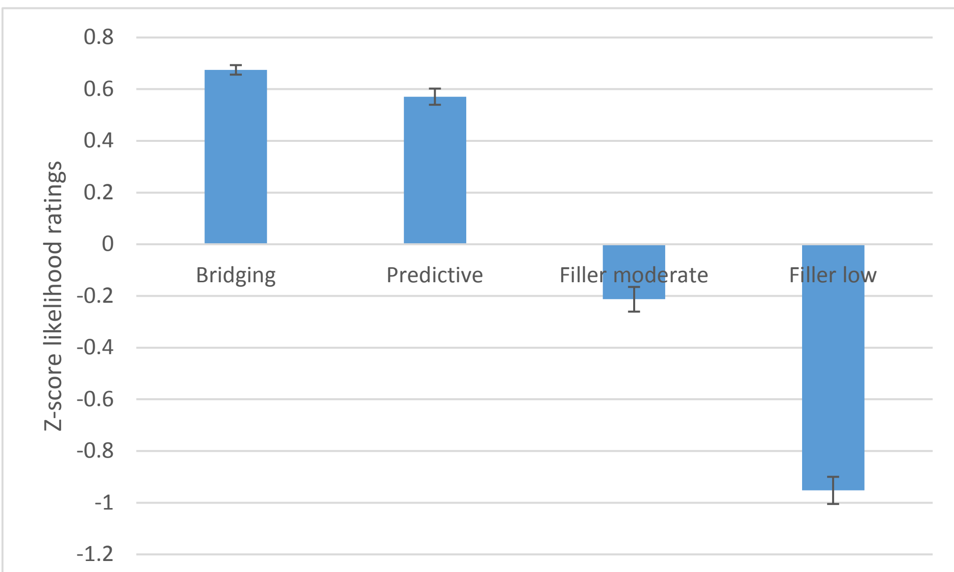

Figure 8. Likelihood Ratings of Inferences

To determine whether bridging and predictive inferences differed when controlling for the semantic associations between the inference and the preceding sentence, I included cosine values from a latent semantic analysis (LSA; Landauer, Foltz & Laham, 1998) in a linear mixed effects regression model focused on bridging and predictive conditions only. LSA values range from -1 to 1, with positive values indicating greater semantic associations. The LSA values for the bridging condition (*M* = .33; *SD* = .12) were lower than those of the predictive condition (*M* = .41; *SD* = .14), $t(60.34) = 2.34$, $p = .022$. Reading comprehension and vocabulary were also included as control variables (they were not included in previous models because the models did not converge). Comprehension and vocabulary did not significantly influence inference ratings, *t*s < 1. When accounting for LSA, readers rated predictive inferences as less likely to occur than bridging inferences, $\beta = -0.07$, $t = -2.56$, $p = .009$ (Table 17). Additionally, in an analysis comparing reading times between sentences preceding bridging and predictive inference ratings, participants spent more time reading sentences that preceded bridging inferences than sentences that preceded predictive inferences, $\beta = -0.29$, $t = -2.86$, $p = .004$, even when accounting for sentence length.

Table 17. Mixed Effects Linear Regressions for Bridging and Predictive Inferences

	β	*df*	*t value*	
(Intercept)	0.73	414	2.92	***
Reading level	-0.02	1537	-2.57	*
Sentence length	-0.01	1535	-2.53	*
Predictive	-0.07	1534	-2.59	**
Reading comprehension	0.00	317	0.55	
Vocabulary	0.01	262	0.86	
LSA	-0.26	1535	-2.33	*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Reading level is calculated as Flesh-Kincaid grade level. Comprehension and Vocabulary measures are from the Gates-MacGinitie AR.

Overall, participants rated bridging and predictive inferences for the final 16 articles as likely to occur. Participants rated bridging inferences as more likely to occur than predictive inferences and read inference-eliciting sentences for bridging inferences more slowly than inference-eliciting sentences for predictive inferences, supporting previous research noting the importance of bridging inferences for comprehension and the optional nature of predictive inferences (Grasser et al., 1994; McKoon & Ratcliff, 1992; Perfetti & Stafura, 2015). Next, an interest norming study was conducted to ensure that these selected articles varied in text interest.

4.1.2 Interest Norming

Across two study sessions, forty-four university students who did not participate in the inference norming study provided interest ratings for the 16 passages normed in the inference study. Twenty-seven were female and 17 were male ranging in age from 18 to 43 (M_{age} = 20.07, *SD* = 5.50). Because the probe-recognition task for Study 3 would be conducted in reading and listening modalities, an auditory version of each newspaper article was created. A female read the articles and recorded them using PRAAT (version 6.0.10, bit version; Boersma & Weenink, 2016) with a 44100 Hz sampling frequency. Articles were recorded in a soundproof booth with a rate of about 145 wpm. Eight counterbalanced lists were created for the newspaper articles, four for reading and four for listening.

Participants were informed that their main task was to evaluate the usefulness of the articles, which was not analyzed in the present study, and provide their thoughts about the articles. After each article, participants provided two interest ratings ("This article was interesting to me" and "I think others would find this article interesting"), a difficulty rating ("This article was difficult to understand"), and evaluated the usefulness of the article (e.g., for an article about baseball, "This article would be useful for writing a paper on pitching techniques"). Ratings were made using a 5-point Likert scale ranging from very untrue to very true.

Participants also completed reading and listening versions of the Gates-MacGinitie comprehension test, the vocabulary portion of the Gates-MacGinitie test, the PPVT-4, the need for cognition questionnaire, reading habits questionnaire, and the reader-based standards of coherence questionnaire. Please see Table 18 for correlations. At the conclusion of the study, participants also completed a questionnaire about modality preferences (listening vs. reading) for the newspaper task and questions about reading strategies was included (Appendix G). This served as a qualitative measure of reading behaviors related to the specific tasks in the experiment.

Of the 44 participants, 39 completed both sessions (providing ratings for all 16 newspaper articles) and 5 completed only Session 1 (providing ratings for half of the newspaper articles).

Table 18. Correlations Among Measures: Interest Norming

	1	2	3	4	5	6	7
1. Gates-MacGinitie Reading	---						
2. Gates-MacGinitie Listening	0.70	---					
3. RB-SOC	-0.01	0.04	---				
4. Need for cognition	0.01	0.12	0.31*	---			
5. Reading habits	-0.08	0.06	0.35*	0.19	---		
6. PPVT-4	0.42*	0.56*	0.32*	0.39*	0.15	---	
7. Gates-MacGinitie Vocabulary	0.37*	0.44*	0.34*	0.22	0.28	0.60*	---

Note. Reading and listening measures are from the Gates-MacGinitie comprehension test (Forms S and T), RB-SOC = Reader-based standards of coherence; PPVT-4 = Peabody Picture Vocabulary Test Revised; * $p < .05$

4.1.2.1 Interest Norming Results

The goal of the study was to examine differences in interest and comprehension difficulty between reading and listening modalities for newspaper articles that would be used as materials in Study 3. Higher difficulty ratings indicate that participants perceived the articles as more difficult and higher interest ratings indicate that participants perceived articles as more interesting. For both reading and listening modalities, higher difficulty ratings negatively correlated with readers' own interest ratings (r = -.53 and r = -.51, ps < .001, respectively). Similarly, for reading and listening, difficulty ratings negatively correlated with readers' perception of others' interest (r = -.47 and r = -.50, ps < .001, respectively). Finally, readers' own interest positively correlated with their perception of others' interest for both reading and listening modalities (r = .65 and r = .64, ps < .001, respectively). Because of the moderate correlation between readers' own interest ratings and perceived interest ratings of others, the scores were averaged to create a composite score.

Next, linear mixed effects regressions tested effects of modality (reading vs. listening) reading skill (a continuous predictor), and their interaction on difficulty and interest ratings. The model included orthogonal contrasts to test differences among the types of articles (thrillers and epidemics vs. all others, thrillers vs. epidemics, laws & treaties and sports vs. all others, laws & treaties vs. sports, and poverty vs. embezzlement and suing). Article nested within subjects was entered as a random slope. Finally, reading grade level, number of words per sentence, number of words in the article, and gender (male vs. female) served as control variables. The linear mixed effects regressions used effects coding, such that the intercept represents the average rating across conditions. Two separate analyses assessed effects on interest and difficulty.

In a linear mixed effects regression, while accounting for article reading grade level, participants rated thriller and epidemic articles more interesting than all other article types (Table 19; Figure 9). Articles about thrillers received higher interest ratings than those about epidemics. Articles about law & treaties and sports received lower interest ratings than all other article types. Articles about law & treaties and sports had similar interest ratings, as did articles about poverty and embezzlement, ts < 1. Participants gave articles in the reading modality higher interest ratings than articles presented in the listening modality. Finally, the modality x reading skill interaction was not significant.

Table 19. Interest Ratings for Newspaper Articles across Various Genres

	β	df	t value	
(Intercept)	3.97	254.70	4.487	***
Reading grade level	-0.14	610.70	-1.809	
Words per sentence	0.18	613.30	3.777	***
Word count	0.00	600.10	-2.264	*
Thriller & Epidemic vs. all	0.96	617.50	5.769	***
Thriller vs. Epidemic	0.55	608.90	2.667	**
History & Sports vs. all	-0.76	612.90	-5.321	***
History vs. Sports	0.08	609.70	0.74	
Poverty vs. Embezzlement and suing	-0.07	608.20	-0.84	
Gender	0.99	39.17	1.417	
Modality	1.14	606.40	2.46	*
Reading Skill	0.28	39.19	1.528	
Modality x Reading Skill	-0.77	606.80	-1.283	

Note. * p < .05, ** *p* < .01, *** *p* < .001. Reading grade level was calculated as Flesch-Kincaid reading level

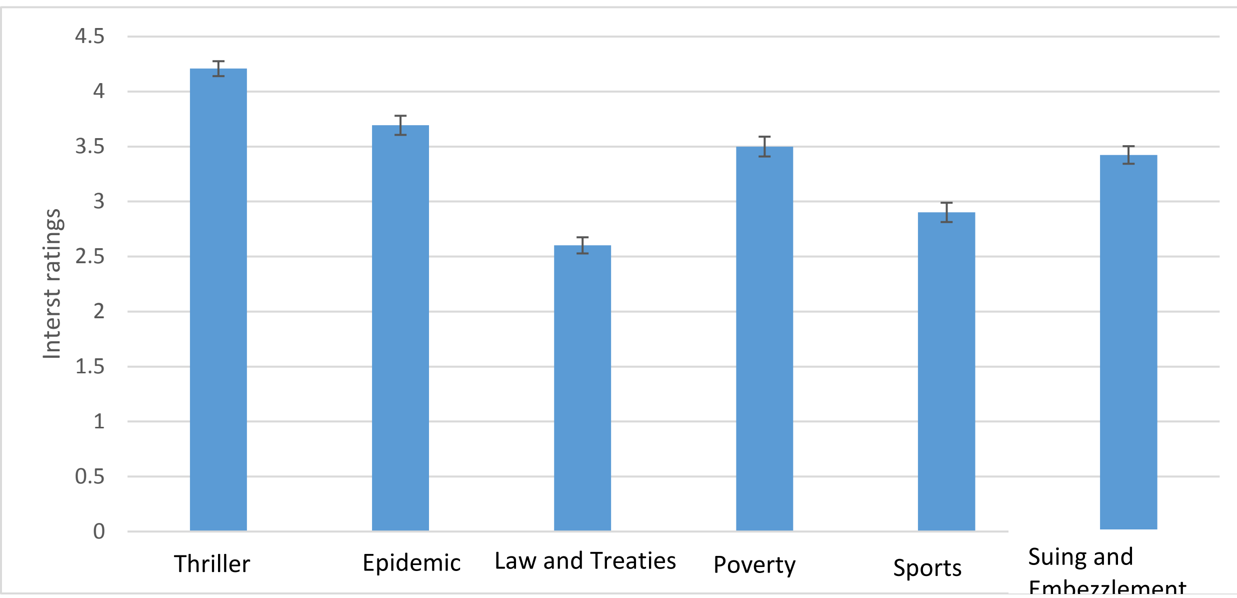

Figure 9. Article Interest Ratings

A similar linear mixed effects regression tested effects on difficulty ratings. While accounting for article reading grade level, participants rated thriller and epidemic articles as less difficult than all other article types (Table 20). Between these two, participants found the thriller articles less difficult than articles about epidemics. Readers found articles about law & treaties and sports more difficult than all other article genres, and articles about law and treaties were rated as more difficult than sports articles. Articles about poverty received similar difficulty ratings to articles about embezzlement and suing. Complementing the interest analysis, participants found it easier to read articles than to listen to them. Higher reading ability marginally predicted lower difficulty ratings, β = -1.06, t= 1.90, p = .065. A significant reading skill x modality interaction, indicated that, as reading skill increased, the difference between reading and listening difficulty ratings decreased (indicating lower ratings of difficulty; Figure 10). Thus, those with lower reading comprehension skills thought articles presented aurally were more difficult to understand than articles they read.

Table 20. Difficulty Ratings for Newspaper Articles Across Various Genres

	β	df	t value	
(Intercept)	2.28	183.7	2.40	*
Article Reading Grade Level	0.04	568.6	0.56	
Words per sentence	-0.09	598.4	-1.81	
Word count	0.00	530.7	2.17	*
Thriller & Epidemic vs. all	-0.43	585.9	-2.58	*
Thriller vs. Epidemic	-0.48	609.3	-2.32	*
History & Sports vs. all	0.73	607	5.08	***
History vs. Sports	-0.25	568.8	-2.34	*
Poverty vs. Embezzlement and suing	0.12	585.3	1.57	
Gender	0.05	40.58	0.21	
Modality	-1.22	408.8	-2.63	**
Reading Skill	-0.95	40.5	-1.13	
Modality x Reading Skill	1.32	414.8	2.20	*

Note. * p < .05, ** p < .01, *** p < .001. Article Reading Grade Level was calculated as Flesch-Kincaid reading level.

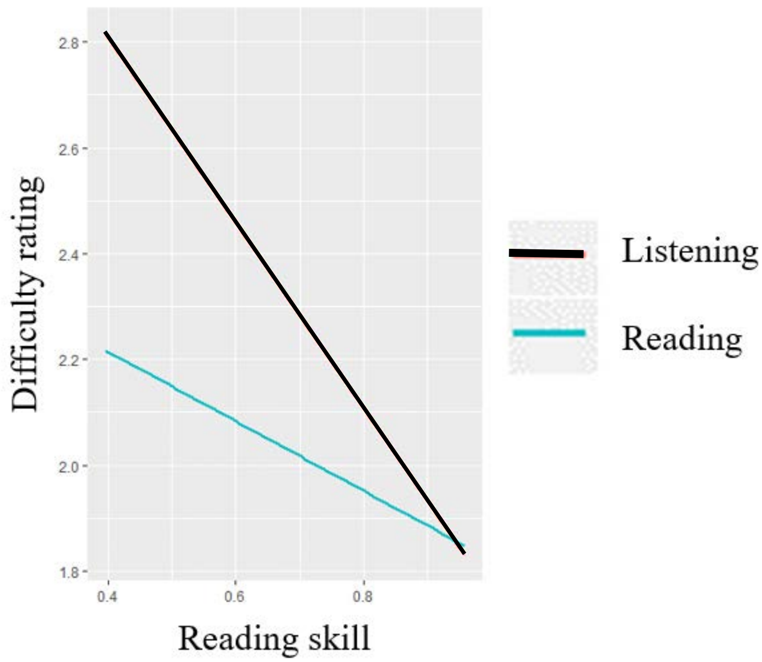

Figure 10. Modality x Reading Skill Interaction

Results from the inference and interest norming studies demonstrate that the materials produced their expected inferences and the articles varied in interest and difficulty. Additionally, participants who rated the articles as more interesting perceived the articles as less difficult.

4.1.3 Participants

The one hundred seventy-eight participants who participated in Study 2 participated in Study 3, which used data from Sessions 1 and 2 (178 participants) and Session 3 (157 participants). The final sample included the 157 participants who voluntarily returned to complete all three sessions.

4.1.4 Probe Recognition Test Materials

Materials in Study 3 included the 16 newspaper articles normed in the inference and interest norming studies to test on-line literal and inferential comprehension. Each participant read and listened to eight passages (counterbalanced across participants). Participants read four passages and heard the other four read aloud. As participants read/heard each passage, they completed a probe recognition task.

At the end of about half of the sentences in the articles, a probe word appeared on the screen in capital letters. Of the probe words, 50% occurred in the preceding sentence and the other 50% did not. Of interest are probes that did not occur in the preceding sentence. These probes were one of four conditions 1) **bridging inferences** required to understand the text, 2) **predictive inferences** related to the preceding sentence but not required for understanding, 3) **paraphrases** of words in the preceding sentence (i.e., literal comprehension), or 4) **control** words unrelated to the preceding sentence. Increased activation of inferences and paraphrases should result in longer response times and reduced accuracy for predictive inferences, bridging inferences, and paraphrases compared to control conditions. Probe words appeared after the end of sentences to limit comprehension disruption. Each article was matched to another article of a similar topic for control words. For example, if “miss” was related to an inference in one article about baseball, it served as a control word in another article about football. See Appendix H for an example newspaper article.

Bridging and predictive inferences were those used in the inference norming task. Paraphrase words had meanings related to a word or phrase in the preceding sentence. For example, “impressive” was considered a paraphrase of “remarkable”. Bridging inferences and their preceding sentences had slightly higher LSA values than those of corresponding control words (in different articles) and their preceding sentence, $t(31) = 2.03$, $p = .051$ (Table 21). Paraphrase and predictive words also had higher LSA values than those of corresponding control words and their preceding sentence. As confirmed by LSA values, paraphrase words were all positively associated with the intended word or phrase in the preceding sentence (*M* = .33, *SD* = .17). Because experimental and control conditions differed in their semantic relatedness, LSA values served as covariates in analyses.

Table 21. LSA Values for Experimental and Control Probes with Their Preceding Sentence

	Experimental condition *M (SD)*	Control condition *M (SD)*	t-value
Bridging	0.33 (.12)	0.28 (.12)	2.03
Predictive	0.39 (1.14)	0.31 (12)	2.71*
Paraphrase	0.34 (.12)	0.28 (.12)	4.60*

Note. * $p < .05$

Each newspaper article contained 24 probes. Each literal, bridging inference, and predictive inference condition contained two probes. Control conditions had six probes and explicit conditions had 12 probes per article. Within the explicit condition, four words occurred in the beginning of the preceding sentence, four words occurred in the middle, and four occurred at the end. The 16 articles had an average word length of 686.94 (*SD* = 78.81), sentence length of 38.93 (*SD* = 4.43), and Flesch-Kincaid grade level of 9.2 (*SD* = 1.66). Across all 16 passages there were 32 probes for literal, bridging inference, and predictive inference conditions for a total of 96 probes. The experimental texts included a total of 96 control probes and 192 explicit probes. The average probe length was 5.95 letters (*SD* = 1.87). Average word frequency (Log HAL frequency) from the English Lexicon Project was 9.61 (*SD* = 1.36) and the range of Log HAL frequency within the corpus is 0-17 (*M* = 6.16, *SD* = 2.40; Balota et al., 2007). Among the explicit, paraphrase, bridging, and predictive inference conditions, no probe word appeared twice. Additionally, probe words did not occur in the article prior to the sentence that preceded them.

4.1.5 Procedure

After completing the Reader-based standards of coherence measure, participants completed reading and listening versions of the on-line probe recognition task (counterbalanced). Participants were randomly assigned to one of 32 randomized lists for article presentation. The lists were constructed from a modified version of a Latin Square, such that no article appeared in the same list as the article it was matched with for the control probe words (e.g., the baseball article never appeared in the same list as the football article).

Participants were randomly assigned to one of two goal conditions: 1) reading/listening for comprehension to verbally answer open-ended questions about the article and 2) reading/listening to complete a phrase matching task. Three question types for the open-ended condition assessed literal comprehension, inferential comprehension, and included an opinion question (e.g., “Do you think the Army or Navy team played better? Why?”). For the phrase matching condition, readers assessed whether three phrases occurred in a previously read article. Across all 16 articles, half of the phrases occurred in an article and half did not. Phrases that did not occur in a specific article came from a matched article on a similar topic, e.g., a phrase that occurred in an article about football was used as a phrase that did not occur in an article about baseball. These phrases were selected to be moderately difficult to match. For example, a phrase that occurred in the article about football but did not occur in the article about baseball was “for it is seldom that such a run is seen in a big game”. Without reading articles, it is plausible that the phrase could occur in either; thus, the task required participants to closely attend to surface-level information of the texts. Phrases had an average length of 10.15 words (*SD* = 3.04).

For the reading modality, the articles were presented one sentence at a time, 2-5 words at a time in black font across a gray computer screen. Words were grouped together in a manner that would not lead to a garden-path sentence. Prior to the start of each sentence, a black fixation cross appeared on the left-hand side of the screen in the location of the first word of the sentence for 500 ms. Then the sentence began. If a probe word appeared after the sentence, a black fixation cross in a white box appeared on the screen for 250 ms, followed by the probe word in capital letters which remained on the screen until participants made a response. The sentence presentation rate was individualized per participant, based on reading times from two shorter newspaper articles read during Session 1. Prior to reading these shorter newspaper articles, participants were assigned their reading goal, answering open-ended questions or phrase matching. The first article was about a woman eloping with her father's driver (170 words, Flesch-Kincaid grade level = 9) and the second article was about steel workers' wages being cut to reduce production costs (276 words, Flesch-Kincaid grade level = 12). These two articles were chosen because they approximate the range of difficulty in the experimental articles. Each participant's reading rate was calculated by averaging time per word across the two articles. A maximum of 500 ms per word was set for the probe recognition task. Participants read a practice article to become familiar with the task.

For the listening modality, the continuous auditory recordings of articles used in the interest norming study were segmented to create 24-25 auditory files per article. Each file represented 1-3 sentences of the article prior to a probe word. An ellipsis appeared in the center of the screen for the duration of each article segment. Similar to the reading modality, a black fixation cross within a white box appeared in the center of the screen for 250 ms prior to capitalized probe words. Probe words remained on the screen until participants made their response.

After each article, participants rated their interest in the article on a scale of 1 (very untrue) to 5 (very true) for the statement "This article was interesting to me". Then, depending on their reading/listening goal condition, participants answered open-ended questions or completed the phrase matching task. Responses to the open-ended questions were recorded on an audio recorder and were not analyzed in the present study.

4.2 Results

The purpose of Study 3 was to examine the influence of reader- and situation-based standards of coherence on reading speed, interest, and on-line comprehension. One hundred seventy-eight participants completed the reading rate task during Session 1 of the study. Of those participants, 85 were in the open-ended condition and 93 were in the phrase matching condition. Of the 157 participants who returned to complete the probe recognition task in Session 3, 72 participants were in the open-ended condition and 85 were in the phrase matching condition Participants remained in the same condition from Session 1 to Session 3. Three participants were removed due to experimenter error and two were removed due to failure to complete the task. Analyses tested the effect of reader- and situation-based standards of coherence on reading speed, followed by their effect on readers' interest. Analyses for the probe recognition task tested the influence of reader- and situation-based standards of coherence on on-line literal (paraphrase) and inferential (bridging and predictive) comprehension.

4.2.1 The Influence of Standards of Coherence on Reading Speed and Interest

Linear mixed effects regressions tested effects of reader-based standards of coherence and reading goal on reading speed and interest for the two short passages participants read during Session 1. Age, gender, and article served as control variables. Additionally, the analyses used the factor weightings from the structural equation model (Study 2) to create composite scores of reading and listening comprehension, decoding, vocabulary knowledge, reading experience, and reader-based standards of coherence. Composite scores were created by computing each participants' Z-score per measure (i.e., the average score for a particular measure was zero across participants) and multiplying the Z-score by the indicator weighting. For example, a participants' Z-score for the PPVT-4 task was multiplied by the indicator weighting of .78. Then, Z-scores for a particular factor were averaged. These composite scores served as control variables in linear mixed effects regression analyses. The composite scores reduced the number of predictor variables and at the same time incorporated the weighted contribution of each task to their hypothesized construct. More importantly, the effects of situation-based standards of coherence (i.e., reading goal) on inference generation can be tested while controlling for reading-specific skills (e.g., decoding), which also affect inference generation.

Results indicate that those in the open-ended condition had longer reading times than those in the phrase matching condition and participants with more reading experience had faster reading times (Table 22). Participants spent more time reading the article about steel workers' wages being reduced (M = 94.90, SD = 33.11), which was more difficult and longer, than the article about a woman eloping (M = 65.65, SD = 24.65). Participants who rated articles as more interesting also had faster reading times. Participants' RB-SOC did not predict reading speed and the RB-SOC x reading goal interaction was not significant.

Table 22. Individual Difference Measures and Reading Goal on Reading Speed

	β	Std. Error	t-value	*p*	
(Intercept)	69.77	11.98	5.82	0.000	***
Age	0.81	0.58	1.39	0.166	
Gender	-2.13	4.30	-0.49	0.622	
Article	25.88	2.59	10.00	0.000	***
Reading Goal	-15.96	4.24	-3.76	0.000	***
RB-SOC	-1.06	6.48	-0.16	0.870	
Decoding	-7.57	6.01	-1.26	0.210	
Vocabulary	5.84	5.41	1.08	0.283	
Listening Comprehension	-5.20	3.37	-1.55	0.124	
Reading Comprehension	3.26	5.28	0.62	0.537	
Reading Experience	-11.68	5.47	-2.14	0.034	*
Interest	-2.93	1.08	-2.71	0.007	**
Article x Reading Goal	0.23	3.32	0.07	0.945	
Article x RB-SOC	-0.41	4.70	-0.09	0.930	
Reading Goal x RB-SOC	6.27	8.54	0.73	0.464	
Article x Reading Goal x RB-SOC	-7.90	6.66	-1.19	0.237	

Note. RB-SOC = Reader-based standards of coherence; Decoding, Vocabulary, Reading comprehension, Listening Comprehension, Reading experience and RB-SOC are composite scores created using factor weightings from the SEM in Study 2. **p* < .05; ***p* < .01; *** *p* < .001.

In a separate analysis, readers with higher RB-SOC and more reading experience rated the articles as more interesting (Table 23). Additionally, readers found the article about a woman eloping more interesting than the article about steel workers' wages being reduced. There was a marginal effect of reading goal on interest ratings $\beta = -.23$, $t = -1.97$, $p = .051$; participants in the open-ended condition rated articles slightly more interesting than participants in the phrase matching condition.

Table 23. Individual Difference Measures and Reaeding Goal on Interest Ratings

	β	Std. Error	t-value	*p*	
(Intercept)	2.68	0.34	7.96	0.000	***
Age	0.02	0.02	1.26	0.208	
Gender	0.19	0.13	1.48	0.142	
Article	-1.04	0.11	-9.78	0.000	***
Reading Goal	-0.23	0.12	-1.97	0.051	
RB-SOC	0.42	0.14	2.98	0.003	**
Decoding	0.17	0.18	0.93	0.353	
Vocabulary	-0.18	0.16	-1.12	0.263	
Listening Comprehension	0.00	0.10	0.01	0.991	
Reading Comprehension	0.10	0.16	0.66	0.507	
Reading Experience	0.32	0.16	1.99	0.048	*
Article x Reading Goal	-0.22	0.21	-1.02	0.310	
Article x RB-SOC	-0.03	0.21	-0.15	0.881	
Reading Goal x RB-SOC	0.27	0.23	1.16	0.246	
Article x Reading Goal x RB-SOC	-0.13	0.43	-0.31	0.758	

Note. RB-SOC = Reader-based standards of coherence; Decoding, Vocabulary, Reading comprehension, Listening Comprehension, Reading experience and RB-SOC are composite scores created using factor weightings from the SEM in Study 2. *$p < .05$; **$p < .01$; *** $p < .001$.

In summary, results indicated that higher reader-based standards of coherence, more reading experience, and, marginally, a reading goal of answering open-ended questions resulted in more interest. Of note, interest and perceived difficulty are correlated (i.e., interest norming study). In the present study, readers with more interest, more reading experience, and a reading goal of phrase matching read the articles more quickly. This supports the rationale for using individualized reading times for the probe recognition task rather than a fixed reading rate for all participants.

4.2.2 Probe recognition Task Results

Results for the probe recognition task in reading and listening modalities are reported separately, because the two tasks differed. In the reading modality, participants read articles and then responded to visually presented probe words, staying within a single modality. In the listening modality participants listened to articles read aloud and then responded to visually presented probe words, switching between listening and reading modalities. Additionally, articles presented in the reading modality were presented to participants based on their own individualized reading rate, whereas all participants heard articles read aloud at the same rate in the listening modality. Therefore, any potential modality effects on the probe recognition task could be because of the aforementioned task differences.

4.2.2.1 Interest Ratings During Probe Recognition Task

Linear mixed effects regressions tested the effect of situation-based standards of coherence, manipulated by reading/listening goal, on interest. Separate analyses for the reading and listening modalities included control variables of age, gender, reading grade level, and word count and composite scores of reading and listening comprehension, decoding, vocabulary knowledge, reading experience, and reader-based standards of coherence. Reading rate was included as a control variable in the reading analysis.

For the reading modality, there was no effect of reading goal or RB-SOC on interest ratings (Table 23). Participants with higher reading comprehension scores had marginally higher interest ratings, $\beta = .30$, $t = 1.82$, $p = .071$. The Reading goal x RB-SOC interaction was not significant.

For the listening modality, participants with higher decoding and RB-SOC rated the articles as more interesting (Table 24). Reading comprehension ability had a marginal effect on interest ratings, $\beta = .28$, $t = 1.77$, $p = .079$. There was no effect of listening goal. A paired-samples t-test compared interest ratings between reading ($M = 2.90$, $SD = 1.22$) and listening modalities ($M = 2.67$, $SD = 1.23$); participants rated articles they read as more interesting than articles they heard, $t(1251) = 3.34$, $p < .001$.

Table 24. Interest Ratings: Probe Recognition Task

		β	Std. Error	t-value	p	
Reading modality						
	(Intercept)	4.77	1.24	3.83	0.001	*
	Reading grade level	-0.14	0.08	-1.80	0.095	
	Word count	0.00	0.00	-0.73	0.477	
	Age	0.01	0.02	0.69	0.491	
	Sex	0.26	0.14	1.85	0.067	
	Reading rate	0.44	0.73	0.61	0.546	
	Decoding	0.22	0.20	1.09	0.278	
	Vocabulary	-0.07	0.17	-0.41	0.685	
	Listening comprehension	-0.07	0.11	-0.68	0.501	
	Reading comprehension	0.30	0.16	1.82	0.071	
	Reading experience	0.21	0.17	1.24	0.217	
	Reading goal	0.00	0.13	-0.02	0.981	
	RB-SOC	0.23	0.15	1.51	0.133	
	Reading goal x RB-SOC	-0.05	0.24	-0.22	0.830	
Listening modality						
	(Intercept)	4.21	1.34	3.15	0.007	**
	Reading grade level	-0.15	0.08	-1.80	0.092	
	Word count	0.00	0.00	-0.58	0.568	
	Age	0.03	0.02	2.10	0.038	*
	Sex	0.27	0.14	1.97	0.051	
	Decoding	0.43	0.19	2.21	0.029	*
	Vocabulary	0.00	0.16	-0.01	0.994	
	Listening comprehension	-0.17	0.11	-1.60	0.112	
	Reading comprehension	0.28	0.16	1.77	0.079	
	Reading experience	0.07	0.16	0.45	0.651	
	Listening goal	0.17	0.12	1.45	0.150	
	RB-SOC	0.31	0.15	2.13	0.035	*
	Reading goal x RB-SOC	0.06	0.24	0.26	0.794	

Note. RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$, ** $p < .01$.

4.2.2.2 Probe Recognition: Reading

Participants performed well on the probe recognition task, with an average accuracy of 90.45% (*SD* = 5.17%). Participants in the open-ended condition scored similarly to those in the phrase matching condition, $t(152) = .21$, $p = .836$ (Table 25). See Table 25 for hit and false alarm rates for each condition.

Table 25. Reading Modality: Probe Recognition Accuracy

Condition	Hit rate *M* (*SD*)	False alarm rate *M* (*SD*)	Overall accuracy *M* (*SD*)
Open-ended	.87 (.08)	.06 (.06)	.91 (.05)
Phrase	.87 (.08)	.06 (.06)	.91 (.06)

Of interest is performance on the probe recognition task for the negative trials on which the probe word did not occur. Analyses used the 96 probe words across the experimental Paraphrase, Bridging, and Predictive conditions along with their corresponding Control conditions. Separate analyses were conducted for reaction time (RT) and accuracy data for each of the experimental and control condition contrasts (i.e., Paraphrase vs. Control; Bridging vs. Control, and Predictive vs. Control). In all analyses, experimental and control conditions contained the same words but had different contexts prior to the probe word. Longer RTs and/or lower accuracy for experimental conditions indicate that readers activated a word related to the probe word, representing literal (i.e., Paraphrase condition) or inferential comprehension (i.e., Bridging and Predictive conditions). Because longer RTs could occur for correct rejections and false alarms, analyses that include both are the primary focus, but for comparison, an additional analysis of only correct rejections was conducted.

Reaction Time Results: Reading

Several data processing steps were taken prior to running analyses. Trials in which probe words had RTs faster than 200 ms were excluded. Additionally, trials in which probe words had RTs greater than 2.5 standard deviations above each participant's average RT across experimental trials were excluded. This resulted in 3.45% of trials being removed. Overall results indicate that participants in both open-ended and phrase conditions had longer RTs and lower accuracy for bridging and paraphrase probes in reading and listening modalities. Differences for predictive vs. control probes depended on readers' interest.

Results are supported by linear mixed effects regressions. All analyses included age, gender, reading rate, reading grade level, sentence length, LSA, and interest as predictors along with composite scores of listening comprehension, reading comprehension, decoding, vocabulary knowledge, reading experience, and RB-SOC (refer to Section 4.2.1 for a description of composite score computation). Analyses also included a RB-SOC x Reading Goal (open-ended vs. phrase matching) interaction to assess the interaction of reader- and situation-based standards of coherence on on-line comprehension. Primary analyses focused on all negative trials in which the probe did not occur with an additional analysis on correct rejections only.

Results for the paraphrase contrast show an overall effect of probe type, in which Paraphrase words had longer RTs than corresponding Control words (Table 26; Figure 11). Additionally, participants with longer reading rates had longer RTs to probe words. No interactions were significant (*p*s > .182). In an analysis of correct rejections only, results were similar, with Paraphrase words having longer RTs than Control words ($\beta = .08$, $t = 5.24$, $p < .001$).

Table 26. Reading RTs: Paraphrase vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	0.49	0.20	2.46	0.016	*
LSA	-0.09	0.08	-1.17	0.242	
Reading grade level	0.01	0.01	1.09	0.295	
Reading rate	0.80	0.30	2.64	0.009	**
Sentence length	0.00	0.00	2.36	0.018	*
Age	0.01	0.01	1.06	0.290	
Sex	-0.01	0.06	-0.12	0.903	

Interest	-0.01	0.01	-1.61	0.108	
Decoding	-0.08	0.08	-0.96	0.340	
Vocabulary	0.13	0.07	1.88	0.062	
Listening Comprehension	0.01	0.05	0.17	0.868	
Reading Comprehension	-0.11	0.07	-1.56	0.122	
Reading experience	-0.13	0.07	-1.82	0.072	
Reading goal	0.03	0.05	0.64	0.526	
Probe Type (Paraphrase vs. Control)	0.08	0.02	5.08	0.000	***
RB-SOC	0.06	0.06	0.96	0.341	
Reading goal x Probe type	0.04	0.03	1.33	0.182	
Reading goal x RB-SOC	-0.09	0.10	-0.88	0.380	
Probe Type x RB-SOC	-0.02	0.03	-0.67	0.501	
Reading goal x Probe type x RB-SOC	0.05	0.06	0.88	0.377	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$; ** $p < .01$; ***$p < .001$.

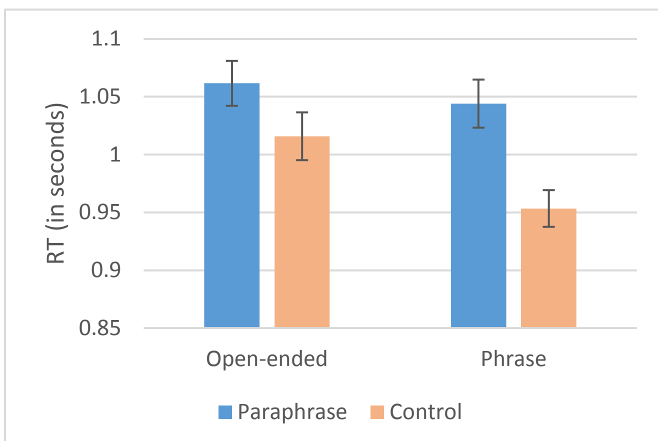

Figure 11. Reading: Paraphrase vs. Control RTs

Results for bridging inferences show a similar pattern to those of the paraphrase contrast. Probe words in the Bridging condition took longer to respond to than those in the Control condition (Table 27; Figure 12). Slower reading rates resulted in longer RTs. No interactions were significant (*p*s > .130). An analysis of correct rejections only showed no significant differences between Bridging and Control conditions (β = -.02, $t = -1.34$, $p = .180$).

Table 27. Reading RTs: Bridging vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	0.52	0.19	2.79	0.006	**
LSA	0.13	0.07	1.85	0.065	
Reading grade level	0.88	0.28	3.10	0.002	**
Reading rate	0.00	0.01	0.29	0.774	
Sentence length	0.00	0.00	-0.14	0.887	
Age	0.01	0.01	1.08	0.284	
Sex	-0.04	0.05	-0.77	0.441	
Interest	-0.02	0.01	-1.91	0.057	
Decoding	-0.05	0.08	-0.63	0.531	
Vocabulary	0.12	0.06	1.93	0.056	
Listening Comprehension	0.01	0.04	0.23	0.817	
Reading Comprehension	-0.11	0.06	-1.78	0.078	
Reading experience	-0.11	0.06	-1.67	0.097	
Reading goal	0.02	0.05	0.40	0.692	
Probe Type (Bridging vs. Control)	-0.04	0.02	-2.33	0.020	*
RB-SOC	0.05	0.06	0.82	0.411	
Reading goal x Probe type	-0.02	0.03	-0.81	0.421	
Reading goal x RB-SOC	-0.04	0.10	-0.42	0.679	
Probe Type x RB-SOC	0.03	0.03	0.91	0.365	
Reading goal x Probe type x RB-SOC	-0.01	0.06	-0.09	0.928	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$, ** $p < .01$.

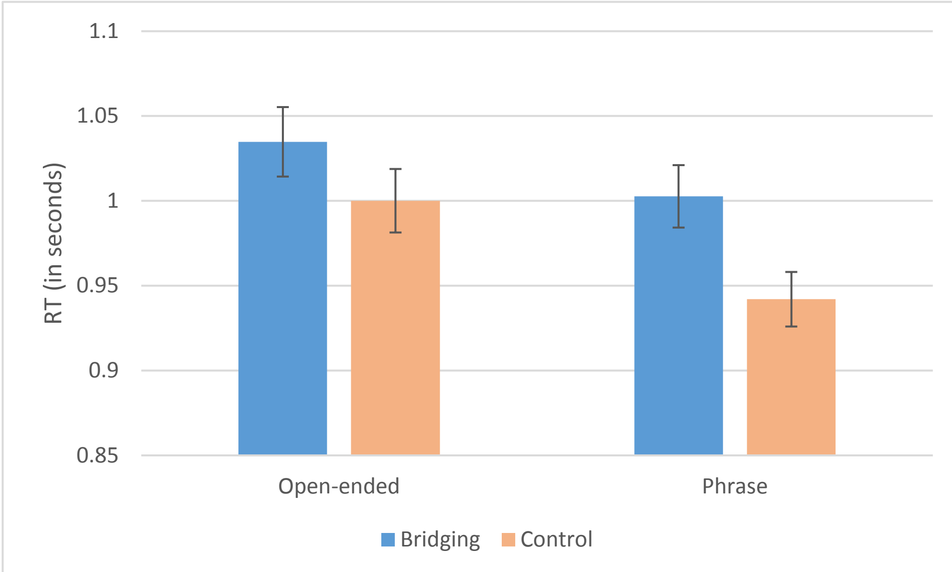

Figure 12. Reading: Bridging vs. Control RTs

Results for predictive inferences were similar; participants took longer to respond to Predictive probes than Control probes (Table 28; Figure 13). Additionally, participants with higher interest ratings responded more quickly. Analyses revealed no significant interactions (*p*s > .372). The analysis for correct rejections only yielded a similar result; Predictive probes had longer RTs than Control probes, β = .05, *t* = 3.08, *p* = .002.

Table 28. Reading RTs: Predictive vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	0.44	0.17	2.65	0.009	**
LSA	0.08	0.06	1.29	0.198	
Reading grade level	0.01	0.01	1.14	0.270	
Reading rate	0.88	0.28	3.17	0.002	**
Sentence length	0.00	0.00	3.56	0.000	***
Age	0.01	0.01	0.84	0.405	
Sex	-0.04	0.05	-0.68	0.495	
Interest	-0.02	0.01	-2.08	0.038	*
Decoding	-0.02	0.07	-0.28	0.782	
Vocabulary	0.08	0.06	1.27	0.206	
Listening Comprehension	-0.01	0.04	-0.18	0.858	

Reading Comprehension	-0.08	0.06	-1.33	0.185	
Reading experience	-0.10	0.06	-1.62	0.108	
Reading goal	0.02	0.05	0.34	0.732	
Probe Type (Predictive vs. Control)	0.05	0.02	3.19	0.001	**
RB-SOC	0.04	0.06	0.67	0.507	
Reading goal x Probe type	-0.02	0.03	-0.76	0.446	
Reading goal x RB-SOC	-0.07	0.10	-0.70	0.486	
Probe Type x RB-SOC	0.00	0.03	0.06	0.954	
Reading goal x Probe type x RB-SOC	0.05	0.06	0.89	0.372	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * *p* < .05; ** *p* < .01; ****p*< .001.

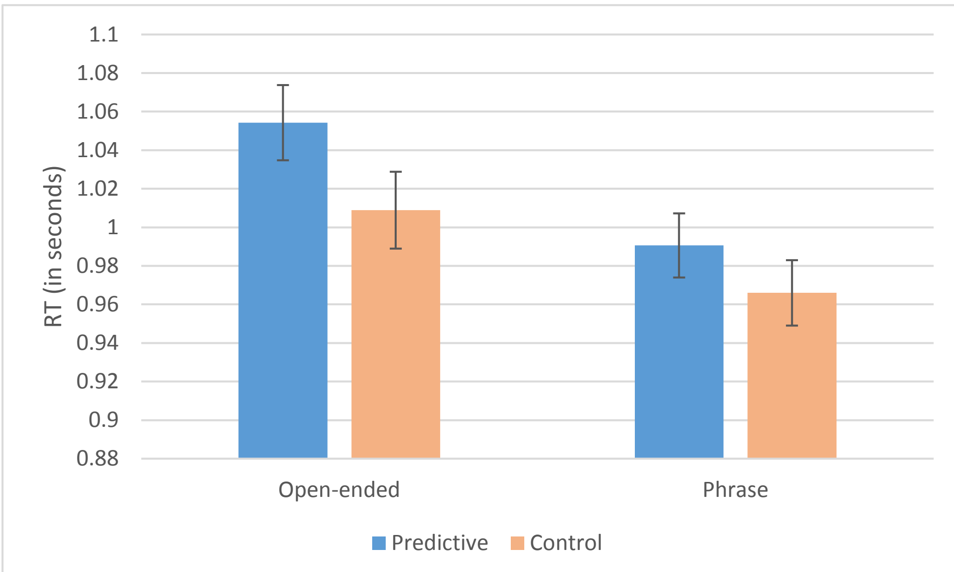

Figure 13. Reading: Predictive vs. Control RTs

As a result of the significant effect of interest, a follow-up analysis examined whether interest interacted with Probe type. Although the difference between Predictive and Control probes was no longer significant, β = -.02, *t* = -.63, *p* = .530, there was a significant Interest x Probe type interaction, β = .03, *t* = 2.22, *p* = .027. As participants' interest ratings increased, Predictive probes had increasingly longer RTs compared to Control probes (Figure 14). Participants with more interest also had faster RTs overall, β = -.02, t = -1.99, p = .047.

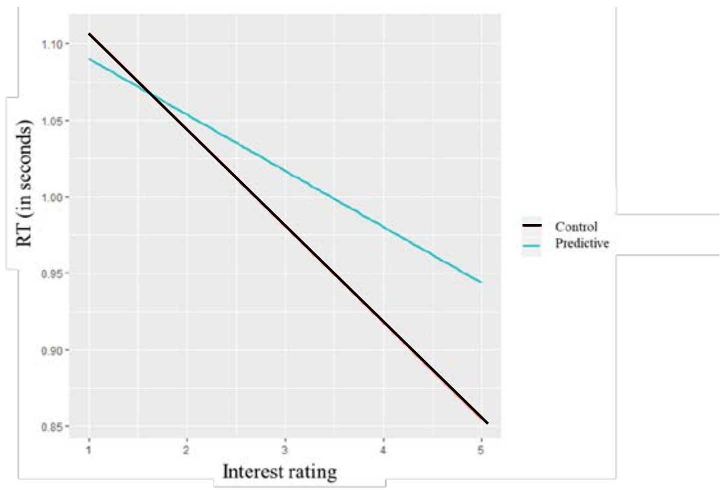

Figure 14. Interest by Probe Type Interaction on RTs

Accuracy Results: Reading

Mixed effects logistic regressions tested effects of probe type on accuracy for Experimental vs. Control words. Participant and article were included as random intercepts. In general, the accuracy data mirror those of the RT data, except probe type did not influence accuracy for Predictive vs. Control words.

Table 29. Reading Accuracy: Paraphrase vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	4.19	0.95	4.40	0.000	***
LSA	-0.07	0.79	-0.09	0.932	
Reading grade level	-0.10	0.06	-1.86	0.063	
Reading rate	0.47	1.29	0.36	0.716	
Sentence length	-0.03	0.01	-2.94	0.003	**
Age	0.00	0.03	-0.15	0.884	
Sex	-0.06	0.25	-0.23	0.820	
Interest	0.41	0.35	1.17	0.244	
Decoding	0.08	0.30	0.26	0.796	

Vocabulary	-0.15	0.20	-0.74	0.459	
Listening Comprehension	0.73	0.29	2.54	0.011	*
Reading Comprehension	-0.14	0.30	-0.47	0.641	
Reading experience	0.17	0.08	2.09	0.037	*
Reading goal	-0.15	0.27	-0.55	0.582	
Probe Type (Paraphrase vs. Control)	-1.66	0.21	-7.72	0.000	***
RB-SOC	-0.39	0.31	-1.27	0.204	
Reading goal x Probe type	0.63	0.41	1.52	0.130	
Reading goal x RB-SOC	-0.52	0.52	-1.00	0.319	
Probe Type x RB-SOC	-0.06	0.42	-0.15	0.885	
Reading goal x Probe type x RB-SOC	-0.66	0.84	-0.78	0.437	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Article difficulty is calculated as Flesch-Kinkaid grade level. **p* < .05; ***p* < .01; ****p* < .001. Subject was included as a random intercept. The model did not converge when Article was included as a random intercept, thus Article was not included.

Mixed effects logistic regressions for accuracy data indicate that Paraphrase (Table 29) and Bridging probe words (Table 30) had lower accuracy scores compared to corresponding Control words. However, accuracy did not differ between Predictive and Control probes (Table 31). No interactions were significant.

Table 30. Reading Accuracy: Bridging vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	4.67	1.07	4.37	0.000	***
LSA	-2.82	0.91	-3.10	0.002	**
Reading grade level	-0.01	0.07	-0.09	0.933	
Reading rate	1.02	1.36	0.75	0.452	
Sentence length	-0.06	0.02	-3.68	0.000	***
Age	0.00	0.03	0.12	0.904	
Sex	-0.12	0.27	-0.47	0.641	
Interest	0.14	0.36	-0.62	0.538	
Decoding	-0.22	0.31	1.99	0.046	*
Vocabulary	0.62	0.22	0.74	0.463	
Listening Comprehension	0.16	0.29	0.84	0.402	
Reading Comprehension	0.24	0.32	-1.61	0.107	
Reading experience	-0.51	0.10	1.40	0.162	
Reading goal	0.11	0.26	0.41	0.685	
Probe Type (Bridging vs. Control)	1.23	0.28	4.32	0.000	***
RB-SOC	0.25	0.31	0.82	0.414	

Reading goal x Probe type	-0.26	0.50	-0.53	0.598	
Reading goal x RB-SOC	0.37	0.53	0.71	0.478	
Probe Type x RB-SOC	-0.09	0.51	-0.17	0.862	
Reading goal x Probe type x RB-SOC	-1.14	1.02	-1.12	0.265	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Article difficulty is calculated as Flesch-Kinkaid grade level. **p* < .05; ***p* < .01; ****p* < .001. Subject and Article were entered as random intercepts.

Table 31. Reading Accuracy: Predictive vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	6.41	1.46	4.38	0.00	***
LSA	-1.45	0.95	-1.53	0.13	
Reading grade level	-0.36	0.11	-3.33	0.00	***
Reading rate	0.52	1.52	0.34	0.73	
Sentence length	0.00	0.02	-0.09	0.93	
Age	0.03	0.04	0.80	0.43	
Sex	0.09	0.30	0.30	0.77	
Interest	0.14	0.11	1.30	0.19	
Decoding	-0.33	0.41	-0.79	0.43	
Vocabulary	-0.12	0.34	-0.36	0.72	
Listening Comprehension	0.25	0.23	1.10	0.27	
Reading Comprehension	0.80	0.35	2.32	0.02	*
Reading experience	-0.33	0.37	-0.90	0.37	
Reading goal	-0.35	0.28	-1.23	0.22	
Probe Type (Predictive vs. Control)	-0.07	0.25	-0.29	0.78	
RB-SOC	0.01	0.33	0.03	0.97	
Reading goal x Probe type	-0.51	0.44	-1.16	0.25	
Reading goal x RB-SOC	-0.34	0.54	-0.62	0.54	
Probe Type x RB-SOC	0.16	0.44	0.35	0.72	
Reading goal x Probe type x RB-SOC	-0.41	0.88	-0.46	0.65	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Article difficulty is calculated as Flesch-Kinkaid grade level. **p* < .05; ***p* < .01; ****p* < .001. Subject and Article were entered as random intercepts.

In summary, RT and accuracy results for Paraphrase and Bridging conditions suggest that readers, regardless of reading goal, activated semantic representations similar to the presented probe words which resulted in longer RTs and lower accuracy scores. For the Predictive inference

RT analysis, increased interest predicted longer RTs for Predictive compared to the Control probes. The accuracy data for the Predictive condition did not show any differences between Predictive and Control probes.

4.2.2.3 Probe Recognition: Listening

Analyses for the probe recognition task in the listening modality were similar to those conducted in the reading modality. Participants had an average accuracy of 87.90% (*SD* = 5.93%) on the probe recognition task. Participants' accuracy did not differ between the open-ended and phrase matching conditions, $t(152) = 1.17$, $p = .249$ (Table 32). See Table 32 for hit sand false alarm rates for each condition.

Table 32. Listening Modality: Probe Recognition Accuracy

Condition	Hit rate *M* (*SD*)	False alarm rate *M* (*SD*)	Overall accuracy *M* (*SD*)
Open-ended	.81 (.09)	.06 (.06)	.87 (.06)
Phrase	.82 (.09)	.06 (.06)	.88 (.05)

Reaction Time Results: Listening

The same data pre-processing steps in the reading analyses were implemented in the listening analyses for negative ("no") trials. This resulted in 3.12% of trials being removed. Initial analyses included correct rejections and false alarms, and an additional analysis focused on correct rejections only. Overall, results indicate that participants in both open-ended and phrase matching conditions were sensitive to bridging and paraphrase probe words, but not predictive probe words.

Results are supported by linear mixed effects regressions. Analyses were nearly identical for those in the reading modality; the differences were that that reading rate and sentence length were not included as control variables.

Results for the paraphrase contrast show that participants took longer to respond to Paraphrase words than Control words (Table 33; Figure 15). Additionally, participants with more interest and higher decoding ability responded more quickly. No interactions were significant (*p*s >.077). An analysis of only correct rejections yielded similar results; participants had faster RTs for Paraphrase probes compared to Control probes, $\beta = .07$, $t = 3.23$, $p < .001$. A follow-up analysis found no interaction of interest with probe type, $\beta = .001$, $t = .07$, $p = .941$.

Table 33. Listening RTs: Paraphrase vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	1.06	0.22	4.90	0.000	***
LSA	-0.05	0.10	-0.51	0.611	
Reading grade level	0.02	0.01	1.93	0.075	
Age	0.00	0.01	-0.19	0.850	
Sex	0.13	0.07	1.81	0.072	
Interest	-0.02	0.01	-2.12	0.035	*
Decoding	-0.23	0.10	-2.15	0.033	*
Vocabulary	0.08	0.09	0.91	0.366	
Listening comprehension	-0.08	0.06	-1.39	0.166	
Reading comprehension	-0.10	0.09	-1.13	0.260	
Reading experience	-0.05	0.09	-0.58	0.562	
Listening goal	-0.04	0.07	-0.62	0.539	
Probe type (Paraphrase vs. Control)	0.08	0.02	3.84	0.000	***
RB-SOC	0.05	0.08	0.65	0.520	
Listening goal x Probe type	0.04	0.04	1.14	0.254	
Listening goal x RB-SOC	-0.17	0.13	-1.32	0.188	
Probe type x RB-SOC	0.04	0.04	1.02	0.308	
Listening goal x Probe type x RB-SOC	-0.02	0.07	-0.28	0.779	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Article difficulty is calculated as Flesch-Kinkaid grade level. * $p < .05$, *** $p < .001$. Subject and Article were entered as random intercepts.

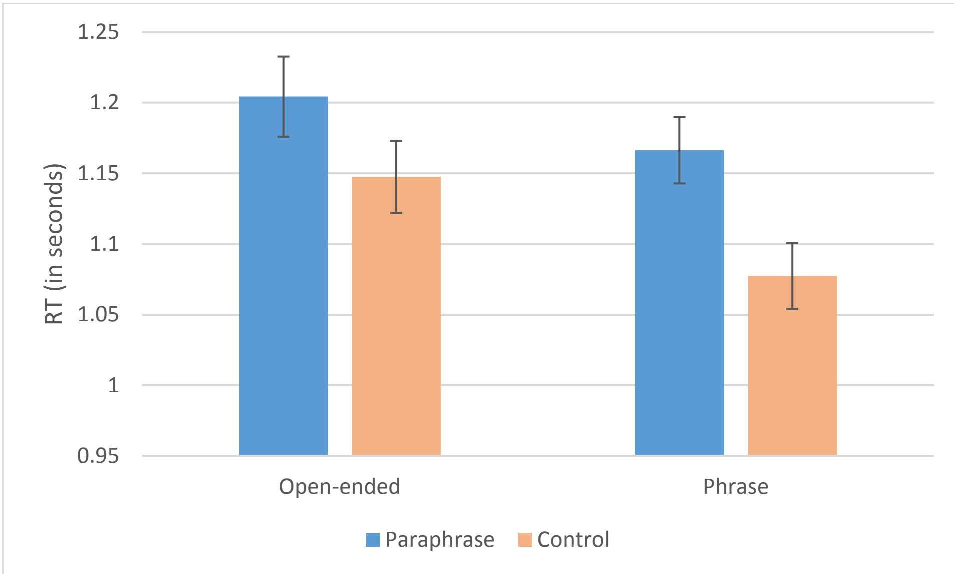

Figure 15. Listening: Paraprhase vs. Control RTs

Results for bridging inferences show a similar pattern; participants had longer RTs for Bridging probes than Control probes (Table 34; Figure 16). Higher LSA values and Interest ratings produced longer RTs. No interactions were significant. In an analysis with correct rejections only, the Bridging condition had longer RTs than the Control condition, β = -.06, t = -3.44, p < .001. A follow-up analysis testing the interaction of Interest ratings and Probe type revealed a non-significant interaction, β = -.02, t = -1.74, p = .240.

Table 34. Listening RTs: Bridging vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	0.89	0.20	4.41	0.000	*
LSA	0.33	0.08	3.91	0.000	***
Reading grade level	0.01	0.01	1.23	0.239	
Age	0.00	0.01	0.36	0.717	
Sex	0.10	0.07	1.48	0.143	
Interest	-0.02	0.01	-2.24	0.025	*
Decoding	-0.18	0.09	-1.93	0.055	
Vocabulary	0.02	0.08	0.24	0.809	
Listening comprehension	-0.04	0.05	-0.83	0.409	
Reading comprehension	-0.06	0.08	-0.78	0.437	
Reading experience	-0.09	0.08	-1.13	0.260	

Listening goal	-0.03	0.06	-0.54	0.593	
Probe type (Bridging vs. Control)	-0.08	0.02	-4.40	0.000	***
RB-SOC	0.06	0.07	0.90	0.368	
Listening goal x Probe type	-0.01	0.04	-0.40	0.692	
Listening goal x RB-SOC	-0.08	0.12	-0.65	0.515	
Probe type x RB-SOC	-0.01	0.04	-0.32	0.747	
Listening goal x Probe type x RB-SOC	-0.08	0.07	-1.10	0.270	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. *$p < .05$, ***$p < .001$. Subject and Article were entered as random intercepts.

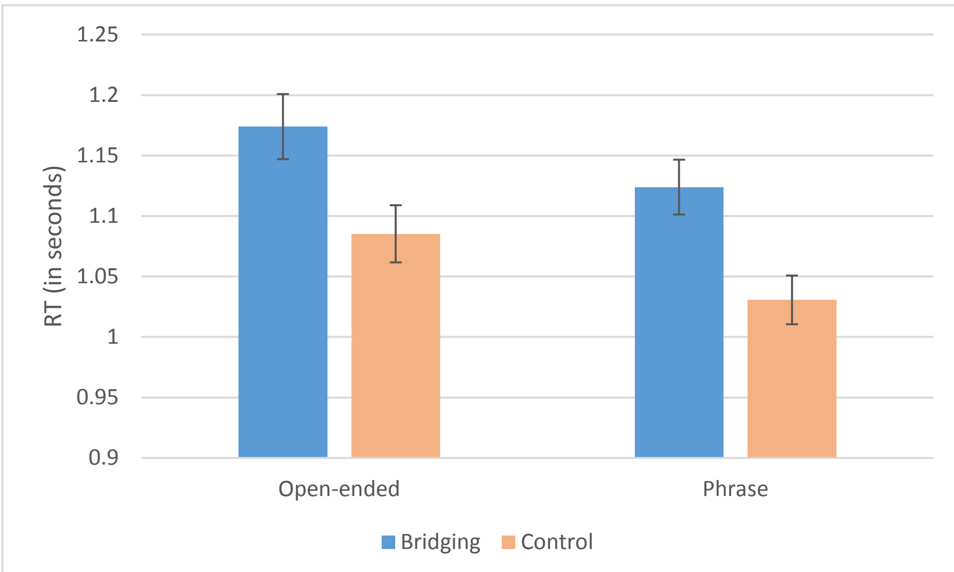

Figure 16. Listening: Bridging vs. Control RTs

Results for the predictive inference analysis revealed no RT differences for Predictive vs. Control probes (Table 35; Figure 17). Increased article difficulty resulted in slower RTs. More interest and higher decoding ability predicted faster RTs. Analyses of correct rejections only produced similar results, with a non-significant effect of Probe type, β = .02, t = -.94, p = .347. There was no interaction of Interest x Probe type, β = .01, t = .437, p = .662.

Table 35. Listening RTs: Predictive vs. Control Probes

	β	Std. Error	t value	p	
(Intercept)	0.96	0.20	4.84	0.000	***
LSA	0.12	0.07	1.60	0.110	
Reading grade level	0.02	0.01	2.18	0.047	*
Age	0.00	0.01	0.06	0.953	
Sex	0.10	0.07	1.39	0.167	
Interest	-0.02	0.01	-2.53	0.012	*
Decoding	-0.21	0.10	-2.12	0.036	*
Vocabulary	0.08	0.08	0.91	0.364	
Listening comprehension	-0.06	0.06	-1.04	0.298	
Reading comprehension	-0.11	0.08	-1.31	0.193	
Reading experience	-0.05	0.09	-0.63	0.528	
Listening goal	-0.03	0.06	-0.53	0.599	
Probe type (Predictive vs. Control)	0.01	0.02	0.81	0.418	
RB-SOC	0.04	0.08	0.52	0.603	
Listening goal x Probe type	0.03	0.03	0.93	0.354	
Listening goal x RB-SOC	-0.15	0.12	-1.17	0.243	
Probe type x RB-SOC	-0.02	0.03	-0.68	0.500	
Listening goal x Probe type x RB-SOC	0.02	0.07	0.33	0.743	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$, *** $p < .001$. Subject and Article were entered as random intercepts.

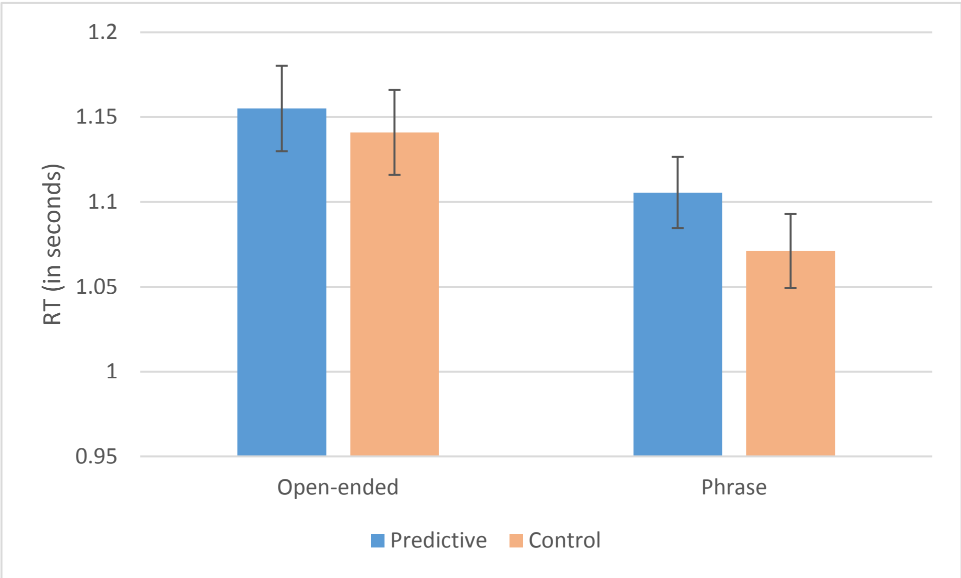

Figure 17. Listening: Predictive vs. Control RTs

Accuracy Results: Listening

In general, the accuracy analyses complement results from the RT analyses. Mixed effects logistic regressions for accuracy data show that Paraphrase (Table 36) and Bridging probe words (Table 37) produced lower accuracy scores compared to corresponding Control words.

Table 36. Listening Accuracy: Paraphrase vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	8.27	0.91	9.13	0.000	***
LSA	-5.28	0.85	-6.21	0.000	***
Reading grade level	-0.45	0.06	-7.93	0.000	***
Age	0.04	0.03	1.29	0.197	
Sex	-0.42	0.20	-2.06	0.040	*
Interest	0.01	0.08	0.19	0.848	
Decoding	0.26	0.29	0.91	0.363	
Vocabulary	-0.11	0.24	-0.45	0.655	
Listening Comprehension	-0.01	0.16	-0.04	0.971	
Reading Comprehension	0.67	0.23	2.95	0.003	**
Reading experience	0.39	0.27	1.45	0.148	
Reading goal	-0.28	0.23	-1.20	0.231	
Probe Type (Paraphrase vs. Control)	-1.43	0.23	-6.19	0.000	***
RB-SOC	-0.44	0.27	-1.60	0.109	
Reading goal x Probe type	-0.09	0.45	-0.19	0.848	
Reading goal x RB-SOC	0.76	0.46	1.64	0.101	
Probe Type x RB-SOC	0.77	0.45	1.72	0.086	
Reading goal x Probe type x RB-SOC	-1.89	0.90	-2.11	0.035	*

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$; ** $p < .01$; *** $p < .001$. Subject and was entered a random intercept. The model did not converge when Article was included as a random intercept.

In the Paraphrase contrast, there was a significant Reading goal x Probe type x RB-SOC interaction (Table 36). For participants in the phrase matching condition, Control probes were answered more accurately than Paraphrase probes, regardless of RB-SOC (Figure 18). Additionally, as RB-SOC increased, accuracy scores also increased. However, for participants in the open-ended condition, as RB-SOC increased, the difference between Control and Paraphrase probes decreased. As RB-SOC increased, participants appeared to be increasingly less accurate for Control probes and more accurate for Paraphrase probes.

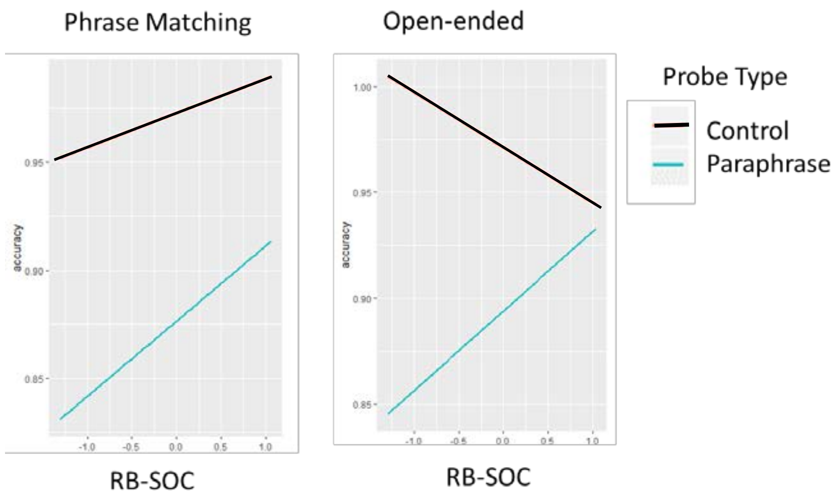

Figure 18. Probe Type x Reading Goal x RB-SOC Interaction

Table 37. Listening Accuracy: Bridging vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	5.83	1.54	3.79	0.000	***
LSA	-4.85	1.01	-4.79	0.000	***
Reading grade level	-0.18	0.13	-1.34	0.180	
Age	0.03	0.04	0.87	0.386	
Sex	0.07	0.29	0.25	0.803	
Interest	0.27	0.11	2.39	0.017	*
Decoding	0.37	0.39	0.95	0.343	
Vocabulary	0.03	0.35	0.09	0.930	
Listening Comprehension	-0.27	0.23	-1.18	0.240	
Reading Comprehension	0.74	0.29	2.54	0.011	*
Reading experience	0.30	0.38	0.80	0.427	
Reading goal	-0.02	0.28	-0.06	0.950	
Probe Type (Bridging vs. Control)	1.00	0.26	3.80	0.000	***
RB-SOC	0.26	0.31	0.83	0.407	
Reading goal x Probe type	-0.03	0.52	-0.05	0.960	
Reading goal x RB-SOC	-0.21	0.51	-0.41	0.685	
Probe Type x RB-SOC	0.79	0.47	1.66	0.096	

Reading goal x Probe type x RB-SOC	-0.17	0.94	-0.18	0.857	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$; *** $p < .001$. Subject and Article were included as random intercepts.

Whereas the RT analyses showed no differences between Predictive and Control probes, the accuracy analysis revealed that participants responded more accurately to Predictive probes than Control probes (Table 38; Figure 19). Across all three analyses the Reading Goal x Probe Type interactions were not significant.

Table 38. Listening Accuracy: Predictive vs. Control Probes

	β	Std. Error	z-value	p	
(Intercept)	8.13	1.91	4.24	0.000	***
LSA	1.43	1.03	1.39	0.164	
Reading grade level	-0.57	0.18	-3.26	0.001	**
Age	0.01	0.03	0.37	0.712	
Sex	-0.06	0.25	-0.23	0.818	
Interest	0.17	0.11	1.66	0.097	
Decoding	0.66	0.34	1.92	0.055	
Vocabulary	0.44	0.29	1.49	0.136	
Listening Comprehension	0.01	0.20	0.07	0.947	
Reading Comprehension	0.01	0.27	0.03	0.978	
Reading experience	-0.05	0.32	-0.16	0.876	
Reading goal	-0.37	0.24	-1.56	0.120	
Probe Type (Predictive vs. Control)	1.12	0.24	4.61	0.000	***
RB-SOC	0.12	0.28	0.43	0.668	
Reading goal x Probe type	-0.41	0.45	-0.93	0.354	
Reading goal x RB-SOC	0.25	0.47	0.53	0.596	
Probe Type x RB-SOC	0.44	0.44	0.99	0.323	
Reading goal x Probe type x RB-SOC	1.29	0.88	1.46	0.146	

Note. LSA = Latent semantic analysis, RB-SOC = Reader-based standards of coherence; Reading grade level is calculated as Flesch-Kinkaid grade level. * $p < .05$; *** $p < .001$. Subject and Article were included as random intercepts.

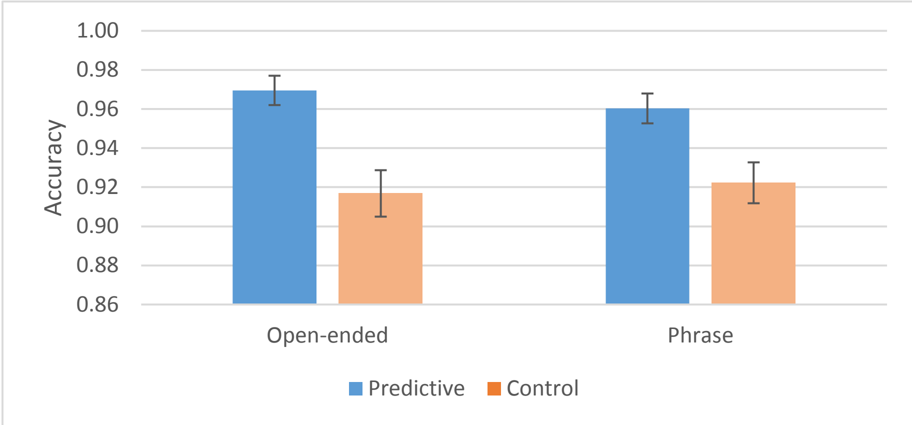

Figure 19. Listening Accuracy: Predictive vs. Control Probes

4.3 Discussion

4.3.1 Summary

Results from the reading rate task demonstrate that participants who read with a goal to answer open-ended questions read more slowly and rated articles as slightly more interesting than those who read the same articles with a phrase matching goal. Readers with more reading experience and readers who rated the articles as more interesting read the articles more quickly than those with less experience and lower interest ratings. Readers with higher reader-based standards of coherence and readers with more reading experience also rated the articles as more interesting. Response time and accuracy results for the probe recognition task indicate that, regardless of comprehension goal, participants had longer RTs and lower accuracy for probes related to paraphrases and bridging inferences than control probes during reading and listening tasks. In the listening modality, participants had similar RTs for probes related to predictive inferences and control probes and participants were more accurate for control probes. In contrast, in the reading modality, as participants' interest ratings increased, they responded more slowly to probes related to predictive inferences than control probes. Finally, participants had similar accuracy scores for probes related to predictive inferences and control probes in the reading modality. Implications of results for the reading rate task are discussed first, followed by those of the probe recognition task.

4.3.2 Reading Speed: A Reading Regulation Strategy

The finding that readers in the open-ended condition read the newspaper articles slower than those in the phrase condition align with previous findings that reading goal influences reading speed (Yeari et al., 2015). Adjusting one's reading speed is one reading regulation strategy that helps readers understand texts. In an eye-tracking study, Yeari et al. presented readers with one of four different reading goals, 1) study for an open-ended test, 2) study for a multiple-choice test, 3) read to give a presentation, and 4) entertainment. Although presentation and study conditions resulted in more re-reading than the entertainment condition, there were also reading time differences between open-ended and multiple-choice study conditions. Participants who read passages to answer open-ended questions spent more time reading passages than those who read the same passages to answer multiple-choice questions. Conversely, a study by Casteel (1993) did not find that reading goal affected reading times for sentences eliciting a bridging inference. In Casteel's study, college students and eighth, fifth, and third graders with a goal to monitor for capital letters had similar reading times as students who had the reading goal to answer questions about events in the story. Because the stories used in Casteel's study were designed to be understood by second graders and adults, it is possible that the texts were too simple to produce reading goal differences. Thus, text difficulty may modulate the effect of reading goal on inference generation.

4.3.3 Influences on Interest

Readers' reader-based standards of coherence—measured via for sub-scales of intrinsic reading goals, extrinsic reading goals and learning strategies, desire to understand and reading regulation strategies, and desired reading difficulty—predicted participants' interest ratings in certain tasks. In the reading rate task, participants read short passages at their own pace with a goal of either answering open-ended questions or phrase matching. Participants with higher reader-based standards of coherence rated articles as more interesting than those with lower reader-based standards of coherence. Similarly, in the listening modality of the probe recognition task, reader-based standards of coherence predicted higher interest ratings, suggesting that reader-based standards of coherence not only reflect one's threshold for understanding written texts, but spoken texts as well. However, in the reading modality of the probe recognition task, readers' reader-based standards of coherence did not predict their interest ratings. One explanation for why reader-based standards of coherence were unrelated to interest ratings in this task is that a certain level of difficulty may be necessary to observe influences of reader-based standards of coherence. Participants' perceived difficulty may have been lower in the probe recognition task in the reading modality than the reading rate task.

In the reading rate task, perhaps reader-based standards of coherence predicted interest ratings because some readers decreased their reading speed to match their reading goal. Longer reading times have been associated with more difficulty (Zabrucky & Ratner, 1992), which is related to lowered interest (e.g., Fulmer & Tulis, 2013). For adults with high reader-based standards of coherence, which relates to enjoyment of reading difficult texts, perceived difficulty may not negatively affect readers' interest ratings. In the reading probe recognition task, reading rates were individualized for participants, eliminating their need to employ a reading regulation strategy of adjusting reading speed. If individualized reading rates reduced readers' perceived difficulty, then perhaps their interest was not reduced, resulting in a non-significant effect of reader-based standards of coherence on readers' interest ratings during the probe recognition task.

Both interest and motivation are correlated with performance on reading comprehension tasks (e.g., Belloni & Jongsma, 1978; McWhaw & Abrami, 2001; Unsworth & McMillan, 2013) and findings from Unsworth and McMillan (2013) suggest that interest may have an indirect effect on comprehension through less mind wandering. In a SEM with university students, Unsworth and McMillan found that interest and motivation for completing a reading comprehension test on a political science text was related to less self-reported mind wandering during reading, and subsequently, better reading comprehension. This result complements claims from other studies that interest may lead to more automatic attention allocation, allowing for more efficient processing (Hidi, 2001; McDaniel, Waddill, Finstad, & Bourg, 2000; Shirey & Reynolds, 1988).

Results from the interest norming study demonstrated that those with lower reading comprehension rated articles they heard more difficult than articles they read; the addition of responding to probe words in the probe recognition experiment most likely increased participants' perceived difficulty, which has been associated with decreased interest (e.g., Fulmer & Tulis, 2013; Horvath, Harleman, & Mckie, 2006). Interest may continually decrease throughout a task. For example, in a study examining the relationship between reading task difficulty and interest with sixth and seventh grade children, Fulmer and Tulis (2013) measured students' topic interest prior to reading a passage above their reading level, twice during reading, and once after reading.

Fulmer and Tulis found that interest decreased as readers progressed through the passages and perceived difficulty predicted lower interest after reading the passages. In the present study, listening to articles that were originally written, completing the probe recognition task, and responding to phrase matching or open-ended question prompts may have been difficult for participants and decreased their interest.

Decoding ability also predicted interest ratings during the probe recognition task in the listening modality and participants gave lower interest ratings to articles they heard read aloud than articles they read themselves. Because decoding ability relies on knowledge of sound-letter correspondences and the ability to transform visual words to verbal code, it should be related to listening comprehension, which involves knowledge of verbal language (Braze et al., 2016; Tunmer & Chapman, 2012). Participants who have better decoding ability may have better sound-letter correspondences and stronger lexical representations, resulting in less effort to activate those representations during the listening task. Less effort for lexical retrieval may result in less overall comprehension effort and more interest. Overall, interest varied by modality, but how did these two factors affect meaning activations ands and inference generation during the probe recognition task?

4.3.4 Influences on On-line Comprehension

The present study used the probe recognition task as a measure of on-line comprehension. Paraphrase words had longer RTs and lower accuracy scores than corresponding control words, suggesting that participants activated relevant semantic representations and encoded them in memory during reading and listening. This result is consistent with several on-line comprehension measures, such as event-related potential paradigms (e.g., Bentin, McCarthy, & Wood, 1985), lexical decision tasks (e.g., Balota & Lorch, 1986; Bentin et al., 1985; Coney, 2002), and naming tasks (e.g., Balota & Lorch, 1986). As mentioned in the Introduction, inhibition effects that result in longer RTs and lower accuracy on probe recognition tasks (Fincher-Kiefer, 1995) can be due to spreading activation from the mental text representation to related words (Anderson, 1983; McNamara, 2005) and/or resonance from the probe word to information in memory (Myers & O'Brien, 1998). In the listening modality, reader-based standards of coherence and comprehension goal interacted with probe type in the accuracy analysis. Participants who had higher reader-based standards of coherence showed smaller accuracy differences between paraphrase and control probes only in the open-ended condition. In the phrase matching condition, participants with higher reader-based standards of coherence were more accurate. The findings suggest that, during listening, participants with higher reader-based standards of coherence were sensitive to different comprehension goals, which affected the memory of semantically related paraphrase words.

As demonstrated by RT and accuracy results, probe words related to bridging inferences had higher levels of activation than corresponding control words in both listening and reading modalities and in open-ended and phrase matching conditions. The finding that bridging inferences were made in both goal conditions reinforces the argument that bridging inferences are necessary for comprehension (Graesser et al., 1994; Potts et al., 1988) and may involve both passive and reader-initiated processes (van den Broek & Helder, 2017). It does not align with previous findings that readers' goals influence the number of bridging inferences they make (e.g., van den Broek et al., 2001). Van den Broek et al. found that participants who read a passage for study made more bridging inferences than those who read the same passage for entertainment. It is possible that, in the present study, having a goal of phrase matching, or any goal that requires remembering text elements, encourages closer reading and deeper levels of understanding than reading for entertainment in which participants do not expect to be tested on the text.

Beyond the RT and accuracy results, the inference norming results demonstrate that bridging inferences are necessary for comprehension and align with previous findings. Similar to a study by Casteel (1993), statements containing bridging inferences received higher likelihood ratings than those containing predictive inferences, even when controlling for sentence length and semantic relatedness between inference statements and inference-eliciting sentences. Reading times for inference-eliciting sentences are also congruent with past findings (e.g., Casteel, 1993; Grasser & Bertus, 1998; Keenan et al., 1984); the inference-eliciting sentences had longer reading times when they required bridging rather than predictive inferences. McKoon and Ratcliff (2015) note that longer reading times alone merely suggest that readers identified a coherence gap, but they do not indicate whether readers resolved this coherence gap. Thus, the combined results of reading times, likelihood ratings, and probe recognition differences (RT and accuracy) demonstrate that readers consistently made necessary bridging inferences that resolved comprehension gaps.

In contrast to bridging inferences, probe words associated with predictive inferences did not consistently have longer RTs or lower accuracy scores than control words. In the reading modality, predictive probes had accuracy scores similar to control probes, in contrast to previous research using the probe-recognition paradigm with predictive inferences (e.g., McKoon & Ratcliff, 1986; Fincher-Kiefer, 1995). The RT results show a different pattern. Participants had increasingly larger RTs for predictive probes compared to control probes as interest increased, indicating that readers with higher interest activated probe words related to a predictive inference to a higher degree than corresponding control words. Participants who expressed interest in what they read may have set higher situation-based standards of coherence and devoted more resources toward making predictive inferences, which enrich a readers' situation model (Grasser et al., 1994; McKoon & Ratcliff, 1986). Longer RTs for probe words related to predictive inferences indicate that probe words may have received increased activation via spreading activation or, through resonance, prompted readers to initiate a context-checking process. Because accuracy was unaffected by probe type, it suggests that the probe words did not receive enough activation to influence response decisions.

Results for bridging and paraphrase control probes for the listening modality mirrored those of the reading modality; bridging and paraphrase probe words had longer RTs and lower accuracy scores than control words. In contrast, predictive probes did not differ from control probes in RT analyses, suggesting that predictive probe words did not receive increased activation from spreading activation or resonance processes. One explanation for no RT effects for predictive probes is that there were limits on resource allocation. Participants constantly switched between listening to articles at a fixed rate and responding to visual words, which may have required more cognitive resources, limiting available resources for making predictive inferences. Interestingly, the accuracy results demonstrate that participants were better at rejecting predictive probe words than control probe words, further supporting the argument that participants did not generate predictive inferences.

In addition to the listening task being more difficult, participants rated articles less interesting when they heard the articles than when they read them. Interest during reading is associated with increased motivation (Deci, 1992; Unsworth & McMillan, 2013), has an indirect relationship with mind wandering (Unsworth & McMillan), and is related to increased use of metacognitive strategies (McWhaw & Abrami, 2001; Pintrich, 1989). If the same principles are applied to listening comprehension, then increased difficulty and decreased interest during listening may result in less motivation, more mind wandering, and use of fewer metacognitive strategies. If making predictive inferences is optional, then when participants are less interested in discourse material and have fewer resources to make predictive inferences, they should make fewer predictive inferences, as seen in the present study.

4.3.5 Interpretations Probe Recognition Tasks

Overall, results suggest that participants consistently experienced inhibition after encountering paraphrase and bridging probes. For predictive probes, this occurred only in the reading modality when participants rated the articles they read as interesting. Slower and less accurate responses could be due to two mechanisms. From the view of spreading activation (Anderson, 1983), participants activated meanings related to paraphrases and generated inferences as they read/heard articles, which would provide evidence that the inferences were made on-line, prior to encountering the probe word. Meaning overlap between activations and probe words would result in longer RTs and lower accuracy. From a memory-based perspective, activations from reading the probe word may have resonated with information in memory, resulting in longer RTs due to context checking only after the probe occurred (Myers & O'Brien, 1998). If the inference was encoded in memory prior to reading the probe word, readers should make more inaccurate decisions, as was seen for paraphrases and bridging inferences. For predictive inferences, it is difficult to disentangle spreading activation and memory-based resonance processes in the present study.

In a series of experiments by Mckoon and Ratcliff (1986), participants read sentences that were followed by a prime, then a probe word that required a "Yes/No" response. When predictive probe words were primed with a word from the context sentence, RTs were longer for predictive probes compared to control probes. Similar results were found when there was no prime, and the probe occurred immediately after sentence offset. However, when an unrelated word (i.e., "ready") served as the prime, RT effects disappeared, suggesting that predictive inferences were only "minimally encoded" and required a cue to provide enough activation to increase RTs. McKoon and Ratcliff (1986) also included an experiment that incorporated a time limit in which participants had to make their response within 600 ms, which they argue does not allow time for context checking.

In the present study, because probe words occurred 250 ms after the end of sentences and participants had longer than 600 ms to respond to the probe, the sentence could have served as a cue, further activating the probe word and prompting the reader to check memory. The probe word itself may have also served as a retrieval cue for text elements activated in memory. Thus, resonance processes or more strategic context checking could have occurred (Forster, 1981; Long et al., 1994; Potts et al., 1988). RT differences between predictive and control probe words occurred for participants who had more interest in the articles they read, but not heard, suggesting that high interest during reading was necessary to increase activation for probe words related to predictive inferences. Because RTs, but not accuracy, were affected by probe type, results suggest that longer RTs for probe words related to predictive inferences were a result of context checking rather than generating a predictive inference prior to reading the probe word. Resonance processes from the probe word to information in memory may have prompted readers to scan their memory for the word, and if the inference was not encoded, then readers found no match in memory and came to an accurate decision that the probe word did not occur. Although readers may not have made predictive inferences prior to the probe word, interest in the articles during reading—and perhaps task difficulty during listening—influenced the activation of predictive inferences.

5.0 General Discussion

The three studies demonstrated varying interrelationships among reader- and situation-based standards of coherence, reading comprehension ability, and comprehension task difficulty on comprehension measures. In Study 1, a measure of reader-based standards of coherence was created and used to assess readers' criteria for understanding texts. The reader-based standards of coherence measure had four sub-scales: intrinsic reading goals, extrinsic reading goals and learning strategies, desire for understanding and reading regulation strategies, and desired reading difficulty. A high criterion for understanding written material was associated with intrinsic motivation for reading and preferring to engage with more difficult reading material. Results from a community sample that included readers with at least a high-school level education showed that readers' extrinsic goals were unrelated to more intrinsic reading factors.

Study 2, which used structural equation modeling, revealed that for college students, extrinsic reading goals were related to the other reader-based standards of coherence sub-scales, suggesting that intrinsic factors and extrinsic reading goals have different relationships for college students compared to adults from a more heterogenous community population. Although the relationships among the four sub-scales of the reader-based standards of coherence measure differed between participants in Study 1 and Study 2, Study 2 replicated the finding that reader-based standards of coherence predicted readers' frequency and duration of reading when controlling for need for cognition, a broader construct of understanding complex problems. Structural equation modeling (Study 2) also provided support for both the Simple View of Reading (Gough & Tunmer, 1986) and the Reading Systems Framework (Perfetti, 1999; Perfetti & Stafura, 2014). Listening comprehension and vocabulary knowledge directly affected reading comprehension, whereas decoding indirectly affected reading comprehension through vocabulary knowledge. Reader-based standards of coherence indirectly affected vocabulary knowledge via reading experience. Thus, there were two routes to vocabulary knowledge, one via decoding ability and the other via reading experience, predicted by readers' reader-based standards of coherence.

Study 3 examined the role of reader- and situation-based standards of coherence on on-line inference generation during reading and listening comprehension tasks. Both reader- and situation-based standards of coherence influenced participants' interest ratings during reading and their interest affected their reading speed. Participants activated relevant semantic representations and made bridging inferences in both reading and listening modalities and in both of the two goal conditions, one directed toward answering open-ended questions, the other deciding on the occurrence of short phrases in the text (phrase matching). Evidence for predictive inferences depended on whether the participants read or listened to the text and also on their rated interest in the text.

The next sections evaluate the contribution of the comprehension assessments used in the present studies for identifying sources of comprehension difficulty. Additionally, implications for the on-line measures on reading interventions are discussed. Finally, the interrelationships among off-line reading-related measures in the SEM and their implications are discussed.

5.1 Evaluation of Reading Comprehension Assessments

Reading comprehension skill was measured by two standardized assessments, the Gates-MacGinitie (MacGinitie et al., 2001) and Nelson-Denny (Brown et al., 1993), which yielded contrasting patterns on literal and inferential questions. Across both tests, lower-skilled readers had a larger difference between literal and inferential comprehension questions compared to skilled readers. However, performance on these two question types depended on the comprehension assessment administered. In Study 2, participants performed better on inferential questions than literal questions on the Gates-MacGinitie test, but the same participants performed better on literal questions on the Nelson-Denny test. One potential explanation for the reversed pattern is that in Study 2 readers could not refer to texts when answering multiple-choice questions for the Gates-MacGinitie test, but they could for the Nelson-Denny test. However, in the inference norming study for Study 3, consistent with the test instructions, participants who took the Gates-MacGinitie test could refer to texts when answering the comprehension questions and participants still performed better on inferential questions.

Whereas the difference between percent correct on the Gates-MacGinitie and Nelson-Denny inference questions was only 3.79%, there was a 21.67% difference between the two tests on literal questions (Study 2). Thus, the main difference seemed to be driven by the literal comprehension questions, with better performance on the Nelson-Denny test than the Gates-MacGinitie test. Literal questions from the Nelson-Denny test often targeted verbatim information from the text and relied on surface-level representations. For example, the text segment, "Momday, who has said that he considers himself primarily a poet,…", could be used to answer the following item, "Momday considers himself primarily…" (answer options: a teacher, a historian, an anthropologist, a poet, a philosopher). Literal questions from the Gates-MacGinitie also included some surface-level questions, but other questions required more vocabulary knowledge and proposition-level representations. For example, the text, "Such occlusions can be as fruitful as they are rare.", could be used to answer the item, "An occlusion of a star by a planet is an event that is…" (answer options: violent, abrupt, uncommon, unpredictable). In summary, some of the Gates-

MacGinitie literal questions required a deeper understanding at the proposition level whereas most of the Nelson-Denny literal questions required surface-level representations. Creating surface-level representations seems to require less effort, especially when readers can refer to texts when answering questions (as in the Nelson-Denny test) and may explain why readers scored higher on the Nelson-Denny literal questions than the Gates-MacGinitie literal questions (Study 2). However, the finding that inference questions were answered more accurately than literal questions on the Gates-MacGinitie test when readers could refer to the text (Study 3) and when they could not refer to the text (Study 2), suggests that proposition-level representations necessary for answering literal questions on the Gates-MacGinitie test were more difficult for readers to create than situation-level representations necessary to answer inferential questions.

The differences in literal questions across the two assessments signifies that the comprehension assessments researchers and educators use may yield different interpretations of readers' comprehension abilities. If the current study included only the Nelson-Denny test, the overall conclusion would be that participants found it easier to answer literal questions compared to inferential questions. However, if the study included only the Gates-MacGinitie test, the conclusion would be that participants performed better on inferential questions. Including both tests had the benefit of assessing surface (Nelson-Denny and Gates-MacGinitie literal questions), proposition (Gates-MacGinitie literal questions), and situation level representations (Nelson-Denny and Gates-MacGinitie comprehension tests).

5.2 Implications for Reading Comprehension Interventions

Many reading comprehension interventions have focused on improving comprehension outcomes for children, with a limited focus on adults (National Reading Report, 2000; National Research Council, 2012). Thus, the present findings from on-line (Study 3) and off-line (Studies 1 and 2) measures of reading comprehension may be useful for informing future reading comprehension interventions that focus on adult learning.

5.2.1 Evidence from On-line Comprehension Measures

The on-line reading rate and probe recognition tasks examined how a participant's comprehension goal (answering open-ended questions or phrase matching) and other factors identified in the SEM (e.g., decoding, reading experience, etc.) related to on-line comprehension processes. Results from the reading rate task demonstrated that readers who had a goal to answer open-ended questions, which required them to engage in close reading and deep understanding, read texts more slowly. In Study 3, these slower reading times may reflect re-reading, an overall slower reading pace, more inferences made during reading, and/or less word skipping. For reading instruction, providing readers prompts that emphasize mastery goals of learning new material should promote high-level text representations (e.g., Ryan et al., 1990), which may result from readers employing reading regulation strategies, such as re-reading difficult sections of the text or asking questions to promote understanding.

In addition to reading goal, readers' interest and reading experiences increased their reading speed. The significant effect of reading experience highlights the importance of reading experience for reading comprehension tasks. This finding does not imply that more experienced readers failed to adjust their reading speed according to the different phrase-matching and question answering reading goals, but instead readers with more reading experience read more quickly relative to those with less reading experience. Thus, more reading experience may be related to better reading fluency—the ability to quickly and accurately read words, phrases, and sentences (Kuhn & Stahl, 2003). Perhaps people with more reading experience—and who engage deeply with texts—had increased familiarity with different syntactic structures, which may have been especially important for reading the newspaper articles that contained unfamiliar syntactic structures from another century (Study 3). Although reading fluency training does not always result in better comprehension (Edmonds et al., 2009; National Reading Panel, 2000), fluent readers may experience less effort and, consequently, more enjoyment during reading. Interestingly, decoding ability did not predict reading speed, suggesting that college readers' ability to use their knowledge of letter-sound correspondences did not affect how quickly they read the texts.

Reader-based standards of coherence also predicted participants' interest in the articles they heard/read in more difficult comprehension situations. In the on-line probe recognition task in the reading modality, readers with more interest in the articles they read had longer RTs for predictive probes compared to control probes. Thus, increasing readers' reader-based standards of coherence and their interest in what they read, especially for struggling readers, may improve readers' predictive inference generation. Interventions that directly encourage predictive inference generation may have less of an effect on predictive inference generation if readers are not interested in the material they read. In a similar vein, tasks that are perceived as more difficult, such as the probe recognition task in the listening modality, may hinder interest. Participants rated passages that they listened to more difficult than passages that they read, possibly because the articles were adapted from their written form and consisted of a structure that is less typical of spoken discourse. Future comprehension interventions should take task difficulty into consideration when measuring inference processes, especially for less-skilled readers. For example, although the probe recognition task is useful for capturing on-line inference generation, task difficulty may reduce the ability to for adults to engage in those processes.

Overall, evidence from the on-line comprehension measures showed that reader-based standards of coherence, interest, and reading experience influenced performance on inference generation tasks and reading regulation strategies while controlling for reading-related skills such as vocabulary knowledge. To evaluate and create reading interventions for adults, it important to identify how these reading-related skills relate to one another and which skills may be targets for improvement.

5.2.2 Evidence from Off-line Comprehension Measures

Study 1 and Study 2 extend the current literature on the relationship between reading motivation and reading comprehension (Davis et al., 2018; Guthrie et al., 2004; Schutte & Malouff, 2007; Wigfield & Guthrie, 1997; Wigfield et al., 2008) by demonstrating that reader-based standards of coherence are indirectly related to reading comprehension. The reader-based standards of coherence measure developed in Study 1 included intrinsic and extrinsic motivations for reading, but importantly included sub-scales measuring readers' desire to understand, reading regulation strategies, and desired reading difficulty. In regression analyses across two samples from different populations, the reader-based standards of coherence measure predicted readers' reading habits. The reading habits questionnaire was a self-report measure, focusing on specific reading experiences as well as typical reading habits such as how often people read a daily newspaper. For the SEM tested in Study 2, reader-based standards of coherence predicted a reading experience latent factor, which included three versions of the author recognition test (Stanovich & West, 1989; Perfetti & Hart, 2002; Acheson et al., 2008), and is a proxy for readers' long-term reading experiences. Thus, the reader-based standards of coherence measure may be a useful tool for assessing the level of text understanding that readers typically aim to have.

Although the SEM results are correlational, findings suggest the following: Readers with higher reader-based standards of coherence read more. Their increased reading experiences leads to increased vocabulary knowledge. Although direct interventions on vocabulary are successful for increasing vocabulary knowledge (Bolger et al., 2008; Kamil, Borman, Dole, Kral, Salinger, & Torgesen, 2008; Mezynski, 1983; National Reading Panel, 2000), these vocabulary gains do not always lead to comprehension gains (Mezynski, 1983; National Reading Panel, 2000). This raises the possibility that an intervention aimed at raising a reader's standards of coherence may lead to more gains in vocabulary and reading comprehension. The next sections review previous findings from interventions that explicitly trained vocabulary knowledge and comprehension strategies, followed by a recommendation for improving standards of coherence, which may encourage adult readers to independently continue building these skills.

5.2.2.1 Vocabulary Interventions

The SEM (Study 2) demonstrated that vocabulary knowledge plays a central role in reading comprehension and, in line with this finding, many interventions with children and adolescents have examined the role of vocabulary knowledge instruction on reading comprehension. Readers can increase vocabulary knowledge via direct instruction or implicitly from spoken or written language (National Reading Panel, 2000; National Reading Council, 2012). The National Reading Panel suggests that instructors pre-teach vocabulary words for specific texts to improve children's comprehension for texts that incorporate newly learned vocabulary words. Learning new words across multiple contexts also facilitates word learning and supplementing contextual word-learning with definitions bolsters this effect (Bolger et al., 2008; National Reading Panel). Higher-skilled readers tend learn new words better than lower-skilled readers (Bolger et al., McKeown, 1985), possibly because higher-skilled readers have richer semantic networks that allow multiple connection-points for newly learned words and/or they have more experience forming new semantic representations. Along with explicit vocabulary instruction across multiple contexts, Kamil et al. (2008) suggest that teachers provide students opportunities to engage in reading, writing, and discussion-based activities that allow them to practice their vocabulary knowledge. Students may also benefit from instruction on how to use context information to learn new words.

For a more general understanding of the effectiveness of vocabulary interventions, Elleman, Lindo, Morphy, and Compton (2009) conducted a meta-analysis on vocabulary interventions for children from pre-K to 12th grade. Out of the 37 interventions conducted from 1950 – 2006, Elleman et al. found that vocabulary instruction increased vocabulary knowledge on experimenter-generated and standardized vocabulary tests. When examining the effect on reading comprehension, comprehension gains occurred for experimenter-generated reading comprehension tests, but not standardized tests. For the experimenter-generated tests, effects of vocabulary instruction on reading comprehension were largest for students with reading disabilities. Similarly, in an intervention with fourth grade students, a SEM found that vocabulary and comprehension instruction for social studies content resulted in increased vocabulary scores and better performance on experimenter-generated materials (Simmons et al., 2010). Simmons et al. found that the interventions also resulted in better performance on social studies reading comprehension tests, but this did not generalize to standardized reading comprehension assessments. Thus, vocabulary instruction may be beneficial for learning specific words, but students' gains in vocabulary knowledge may not transfer to standardized reading assessments that do not include trained words (Kamil et al., 2008).

5.2.2.2 Comprehension Strategy Interventions

Beyond needing adequate vocabulary knowledge for text comprehension, readers must also accurately use comprehension strategies. The reader-based standards of coherence measure (Studies 1 and 2) included items that assessed reading strategies, such as re-reading, summarizing, outlining, and using context to learn new words (Appendix D). Reader-based standards of coherence's relationship to reading experience suggests that readers who report using reading comprehension strategies tend to have more reading experiences, which relates to higher vocabulary knowledge. The National Reading Panel (2000) recommends that teachers facilitate comprehension processes via comprehension strategies, such as teaching story structure, summarization, question generation, comprehension monitoring, and activating relevant background knowledge. Teaching text structures, such as compare-contrast structures, results in improved reading comprehension (Williams, 2005), especially for poor readers (National Reading Panel).

Although strategy instruction is recognized as a method to increase reading comprehension (Kamil et al., 2008; National Reading Panel, 2000; National Research Council, 2012; McKeown, Beck, & Blake; 2009), strategy instruction may not always generalize to increased performance in novel contexts, such as standardized reading comprehension tests (e.g., Edmonds et al., 2009). In an intervention with fifth grade students, McKeown et al. tested whether strategy instruction (summarizing, predicting, generating inferences, and comprehension monitoring) was better than answering open-ended questions for gains in reading comprehension. In this intervention, students received scaffolded instruction from their teachers and engaged in group discussions. Students in the strategy instruction group were given names and definitions of strategies along with discussions of how the strategies could be implemented. Students in the open-ended question instruction group received modeling of how to ask and answer open-ended questions useful for understanding the content of texts. Students' recall of text information was better in the open-ended instruction group than the strategy instruction group; however, in a transfer task of silent reading, the two groups did not differ on recall. Furthermore, in the silent reading task, a control group that included discussion, but not instruction, on strategies or answering open-ended questions performed similarly to the strategy and open-ended instruction groups. McKeown et al. suggest that encouraging young readers to be actively engaged via discussion may have equally benefited all three groups. Other research also suggests that encouraging students to be actively engaged readers may lead to comprehension benefits (Kamil et al., 2008; Gernsten, Fuchs, Williams, & Baker, 2001). More specifically, increased engagement may result in higher standards of coherence, which should in turn increase reading comprehension performance.

5.2.2.3 Reader-based Standards of Coherence and Effective Reading Experiences

Vocabulary and comprehension strategy interventions have demonstrated that readers can improve these skills, which relates to comprehension gains on material of related content. The SEM (Study 2) suggests that readers who have high reader-based standards of coherence use more reading comprehension strategies and have more reading experience, which is associated with greater vocabulary knowledge. In an effort toward implementing reading comprehension interventions that not only train specific skills but equip students to continue to develop these skills independently, future interventions should focus on ways of increasing reader-based standards of coherence.

One avenue to improve readers' standards of coherence is via increasing their motivation and engagement during reading. Research from the motivation literature demonstrated that students with mastery goals, i.e., goals to increase performance and ability in a specified domain (Ames & Archer, 1988), were more likely to continue pursuing that activity in the future (Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000). Furthermore, college readers with high intrinsic motivation were more likely to remember more text elements (Ryan, Connell, & Plant, 1990).

In interventions with children that focus on reading strategies and reading motivation, readers with increased motivation had higher reading comprehension outcomes (e.g., Wigfield & Gutrie, 1997; Wigfield et al., 2008) and readers who found utility in what they read were more likely to have greater comprehension than those who did not (McWhaw & Abrami, 2001). In a 12-week intervention on comprehension strategies and reading engagement, Wigfield et al. (2008) found that fourth grade readers who received training for reading motivation strategies along with comprehension strategies reported higher reading engagement at the end of the intervention. The motivation component included providing students with content goals, giving students autonomy, hands-on activities, using interesting texts during instruction, and collaboration activities. The comprehension strategies component included activating relevant background knowledge, summarizing, questioning, organizing graphically, and identifying story structures. Readers who received both types of training performed better on standardized and experimenter-generated tests than those who received only comprehension strategy training. Importantly, reading engagement mediated the effect of the motivation and comprehension strategies intervention; readers who experienced greater increases in reading engagement showed greater gains on reading comprehension measures.

When readers are engaged, they may set higher standards of coherence and seek out more reading experiences on their own. The SEM results suggest that these increased reading experiences are related to higher vocabulary knowledge, which influences reading comprehension. Of note, more reading experience may not always result in reading comprehension gains. A review of reading comprehension interventions by the National Reading Panel (2000) indicated that interventions on sustained silent reading, a practice often employed in schools, have produced little support that simply reading more results in reading comprehension gains. During sustained silent reading, children often engage in unguided reading for a specified amount of time, with little to no discussion or comprehension questions about what they read. Many of these interventions did not closely monitor whether children actually engaged in silent reading or if their overall reading experiences changed, making it difficult to accurately assess the influence of readers' silent reading experiences on reading comprehension (National Reading Panel).

Although more controlled experiments on silent reading need to be conducted, research indicates that supplying readers with relevant, meaningful reading prompts may promote engagement (National Reading Council, 2012). In recognizing the importance relevant reading situations, Sabatini, O'Reilly, Halderman, and Bruce (2014) created a standardized reading assessment for children called the global, integrated scenario-based assessment that provides students with real-world scenarios, such as looking up information to create a website about farming. Providing readers with effective reading experiences that promote deep levels of understanding may be beneficial for vocabulary and comprehension gains. Integrating standards of coherence with reading is one way to create effective reading experiences. Higher standards may result in close reading of texts, which is associated with extracting meaning beyond what is explicitly written (Partnership for Assessment of Readiness for College and Careers (PARCC);

2011). Close reading involves the use of background knowledge and knowledge about text structures to come to multiple implicit meanings of a text. During close reading, readers find evidence for a particular interpretation of the text and evaluate the strength of that evidence through their own understanding, the wording of the text, and what they believe the author is trying to convey (Boyles & Scherer, 2012; Fisher & Frey, 2012; PAARC, 2011). Only at high standards of coherence can readers employ such deep reading processes.

Based on findings from the present study along with findings from previous reading comprehension interventions (mostly with children), it may be beneficial for future reading comprehension interventions with adults to integrate skill instruction with ways to increase reader-based standards of coherence. Evidence suggests that adult readers can learn new words from context (e.g., Bolger et al., 2008); however, more reading experience alone may not result in vocabulary gains (National Reading Panel, 2000). Readers should be motivated to learn from what they read, setting high standards of coherence that allow for deep, close reading during reading experiences. Finally, although the SEM (Study 2) did not include all possible individual difference measures, such as working memory or general intelligence, it provided a view of the unique interrelationships among reading-related skills and reader-based standards of coherence.

5.3 Conclusions

In summary, the present studies demonstrated that a self-report measure of reader-based standards of coherence indirectly influences reading comprehension by increasing reading experiences. The interrelationships among reader-based standards of coherence factors may depend on how a reader internalizes extrinsic reading goals. The on-line study shows that comprehension skill alone does not account for on-line processes, but that factors such as text interest, reading experience, and reader-based standards of coherence are important for understanding the types of inferences that readers make.

Appendix A

A1. Need for Cognition (Cacioppo et al., 1984)

1. I would prefer complex to simple problems.
2. I like to have the responsibility of handling a situation that requires a lot of thinking.
3. Thinking is not my idea of fun. *
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. *
5. I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something. *
6. I find satisfaction in deliberating hard and for long hours.
7. I only think as hard as I have to. *
8. I prefer to think about small, daily projects to long-term ones. *
9. I like tasks that require little thought once I’ve learned them. *
10. The idea of relying on thought to make my way to the top appeals to me.
11. I really enjoy a task that involves coming up with new solutions to problems.
12. Learning new ways to think doesn’t excite me very much. *
13. I prefer my life to be filled with puzzles that I must solve.
14. The notion of thinking abstractly is appealing to me.
15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort. *
17. It’s enough for me that something gets the job done; I don’t care how or why it works. *
18. I usually end up deliberating about issues even when they do not affect me personally.

* Items are reverse scored

Appendix B

B1. Adult Reading History Questionnaire

Finucci et al., 1982, 1984; Lefly & Pennington, 2000; Parault & Williams, 2009

Table 39. Items Included in the Adult Reading History Questionnaire

Item	Response type	Response options	Details	Citation
How much reading do you do for pleasure?	Likert	None - a great deal	6-point scale with halves (3 - 0)	Lefly & Pennington (200)
How much reading do you do in conjunction with your work? (If you are a full-time student, you can consider that your job)	Likert	A great deal - none	6-point scale with halves (3 - 0)	Finucci et al. (1982)
Do you read daily (Monday-Friday) newspapers	Likert	Every day; once a week; once in a while; rarely; never	4-point scale with halves (3 - 0)	Finucci et al. (1984)
Do you read a newspaper on Sunday?	Likert	Completely every Sunday; scan each week; once in a while; rarely; never	4-point scale with halves (3 - 0)	Finucci et al. (1984)
Do you read newspaper and/or magazine articles on the internet?	Likert	Every day; once a week; once in a while; rarely; never	4-point scale with halves (3 - 0)	Finucci et al. (1984)
How many books do you read for pleasure each year?	MC	None; 1-2; 2-5; 6-10; More than 10		Finucci et al. (1984)
Please estimate the number of hours you read last week.	MC	Less than 5; 5 - 10; 11-20; 21-30; more than 30		Parault & Williams (2009)
How many magazines do you read for pleasure each month?	MC	5 or more; 3-4 regularly; 1-2 regularly; 1-2 irregularly; None		Finucci et al. (1984)

Note. MC = Multiple-choice.

Appendix C

C1. Reader-based Standards of Coherence Measure: 87 items

Table 40. Full Orignal RB-SOC Measure

Number	Item	Intrinsic Reading Goals	Extrinsic Reading Goals and Learning Strategies	Desire to Understand and Reading Regulation Strategies	Desired Reading Difficulty
1	I find satisfaction in finishing a difficult text.	0.31	0.12	0.19	0.11
2	I try to keep track of characters' goals and motives when I read.	0.14	0.26	0.35	0.06
3	When I read I usually skip words that I do not know.	-0.05	-0.11	0.32	0.39
4	I read to learn more about a topic I'm unfamiliar with.	-0.13	0.09	**0.68**	0.01
5	If I really want to comprehend a book/article, I try to put myself in an environment that will help me do so (e.g., find a quiet place).	0.01	0.09	**0.54**	-0.14
6	When I read for something important, I tend to underline or highlight sections, titles, concepts, and/or examples.	0.12	0.49	-0.07	0
7	I commonly read to keep up with current events.	-0.15	0.29	0.37	0.04
8	I'm usually more concerned about finishing a text than comprehending it.	-0.08	-0.28	**0.58**	0.28
9	I often wish I had more time to read for pleasure.	**0.64**	0.01	0.11	-0.08
10	I often skim texts when I read for work/school.	-0.18	0.03	0.15	**0.51**
11	I often try to understand the author's point of view and his/her arguments.	0.09	0.22	0.49	0.08
12	One benefit of reading is that it can improve my vocabulary.	0.33	0.09	**0.51**	-0.02
13	In comparison to other activities I do, it is not very important for me to read a lot.	0.42	-0.06	0.2	0.3

14	I am confident I can understand difficult books or articles.	0.25	0.16	0.29	0.18
15	If a text is difficult, I often re-read sections of the text.	0.14	-0.05	**0.67**	-0.23
16	I read to improve my work or university performance.	0	**0.53**	0.2	0
17	When I'm reading for fun, I rarely look up a word I do not recognize.	0.01	0.05	0.22	**0.51**
18	It is important for me to read for fun, despite my other obligations.	**0.74**	0.05	-0.02	0.13
19	To tell whether I really understand something, I try explaining to myself what the author is trying to convey.	0.04	**0.51**	0.19	0.01
20	When I read, I frequently check to make sure I'm understanding the text.	0	0.39	0.36	-0.08
21	Once I've started reading a story or article, I continue reading to find out its conclusion.	0.17	0.06	0.33	0.12
22	I'd rather read easy texts than difficult ones.	0.04	0.14	-0.07	**0.75**
23	I often think about what I've read even while engaging in other activities.	0.38	0.12	0.2	-0.02
24	When I don't know words in a text, I often look them up.	0.07	0.18	0.4	0.25
25	I most often read to relax.	**0.8**	0.07	-0.24	0.01
26	I routinely make outlines to help me understand what I read.	-0.01	**0.72**	-0.23	0.02
27	When I read, I typically take note of key words or concepts in each paragraph.	-0.1	**0.67**	0.05	0.05
28	I generally enjoy explaining what I've read to a peer or co-worker.	0.33	0.44	0.1	0.03
29	I usually only read when I have a lot of free time, like on vacations or long commutes.	0.02	-0.2	0.12	0.4
30	I'd rather read a summary than spend effort understanding a long text.	0.31	-0.16	0.06	**0.59**
31	I typically learn new concepts through reading.	0.1	0.17	**0.56**	0.11
32	Reading helps me improve my work performance.	0.12	**0.51**	0.22	0.08

33	Reading is a necessary part of my job.	0	0.41	0.19	0.03
34	I like to discuss what I've read with others.	0.48	0.26	0.1	-0.1
35	When I read, it is usually because I must for school/work.	0.39	-0.4	0.08	0.27
36	I often make summaries of what I read.	-0.08	**0.67**	-0.19	0.03
37	If I'm not reading for work/school, I would rather be entertained by a text than informed by it.	-0.4	0.13	0.14	0.45
38	I read more texts outside of school/work than most of my peers.	0.43	0.23	-0.02	0.28
39	I compare different points of view when I read.	-0.03	**0.54**	0.18	0.16
40	When I don't know words in a text, I try to figure out the meaning based on surrounding information.	0.06	0.03	**0.56**	-0.2
41	When I read, I generally try to see how sentences are related to one another.	0.09	0.45	0.24	-0.03
42	I read so I can understand how other people think.	0.06	0.49	0.2	0.09
43	I read so I can have informed conversations with peers/co-workers.	-0.06	**0.64**	0.13	-0.01
44	I tend to read more quickly when I read a difficult text.	-0.17	-0.34	0.48	0.21
45	One reason I like to read is that it allows me to exercise my imagination and creativity.	**0.68**	0.01	0.14	0.06
46	I read to study for exams or a job promotion.	-0.15	**0.57**	0.01	-0.1
47	One reason I read is to improve my reasoning skills.	0.08	**0.61**	0.14	0.11
48	I often skim when I do not need to know the content very well.	-0.01	0.02	0.09	**0.58**

49	When I read to find an answer to a specific question it's enough for me that I find the right answer; I don't care about the reason behind the answer.	-0.09	-0.02	0.44	0.43
50	I try to really understand a text, so I can form my own opinions about the content of the text.	0.14	0.13	**0.69**	0.03
51	I read to help me fall asleep.	-0.33	-0.08	0.38	0.05
52	I gain new insight on how to solve problems by reading.	0.05	0.33	0.47	0.1
53	When I'm having trouble understanding something, I read that section of the text more quickly.	-0.14	-0.37	0.47	0.37
54	Increasing my vocabulary knowledge through reading is important to me.	0.35	0.36	0.15	0.17
55	If a text does not explain a concept clearly enough, I'll look it up somewhere else.	0.09	-0.07	0.48	-0.03
56	I tend to re-read sections of a text more slowly/carefully than when I read them the first time.	-0.12	0.22	0.27	-0.03
57	I think it is important to be a good reader for my future/current career.	0.24	0.39	0.16	-0.08
58	I often look up concepts I do not know.	0.04	0.19	**0.62**	0.11
59	I generally do not like reading dense texts.	0.15	0.16	-0.16	**0.73**
60	I like learning new information when I read.	0.18	0.06	**0.6**	0.03
61	I read so I can show others that I'm knowledgeable about a topic.	0.03	**0.56**	0.07	-0.05
62	I get frustrated when a text does not make sense.	0.05	-0.12	0.22	**0.54**
63	Reading is not my idea of fun.	**0.62**	-0.2	0.06	0.32
64	Understanding what I read is part of the fun of reading.	**0.53**	0.03	0.34	0.07
65	I often highlight or underline words to follow an author's line of thinking.	0.03	**0.7**	-0.15	0.07

66	For a lot of my reading, I'm not concerned with how well I understand.	-0.02	-0.25	**0.51**	0.36
67	I have a strong desire to understand what I'm reading when I read for fun.	0.4	-0.02	0.4	0.06
68	If I cannot guess what a word means based on its context, I will look it up.	0.21	0.04	0.33	0.2
69	I tend to not read texts if there is an option to watch a movie or video on the topic instead.	0.45	-0.18	0.06	**0.53**
70	I generally pay attention to details as much as general themes when I read.	0.22	0.19	0.41	0.07
71	I normally relate what I read to my life.	0.37	0.28	0.18	0.03
72	When I start reading a text, I feel obligated to finish it.	0.21	0.13	0.14	0.1
73	I frequently read to learn new information about topics that interest me.	0.06	0.12	**0.54**	0.18
74	The texts I read for leisure have had a significant impact on my way of thinking.	0.32	0.24	0.19	0.09
75	I read to improve my day-to-day life.	0.39	0.42	0.06	0.06
76	For difficult concepts/texts, I try to visualize what is going on.	0.26	0.1	0.48	-0.14
77	I like to read texts that don’t require much effort to figure out.	0.04	0.17	0.04	**0.6**
78	I’d rather read short texts than long ones.	0.27	0.04	-0.13	**0.57**
79	When I read, I nearly always want to understand what I've read.	0.3	-0.11	**0.63**	-0.18
80	I read because I enjoy it.	**0.79**	-0.07	0.15	0.03
81	I re-read sections of the text if something does not make sense.	0.16	-0.11	**0.64**	-0.08
82	I try to sound out words that I don’t know.	0.12	0.2	0.27	-0.01
83	I typically read at the same speed, regardless of my reading purpose.	-0.23	-0.12	0.05	-0.05

84	I feel that it is important to read regularly.	**0.69**	0.02	0.19	0.09
85	I usually read to learn new skills (e.g., learn a recipe, how to build something).	0.01	0.35	0.23	-0.14
86	When I read stories, I like predicting what happens next.	0.36	0.27	0.02	0.03
87	I try to avoid situations where there is a high chance I will have to read a text carefully.	0.28	-0.21	0.17	**0.51**

Note. Items in bold have factor loadings > .5.

Appendix D

D1. Finalized Reader-based Standards of Coherence Measure: 31 items

Prompt: This portion of the questionnaire focuses on your thoughts about reading. For each of the statements, please indicate to what extent the statement is characteristic of you. If the statement is extremely uncharacteristic of you please select "Strong disagreement"; if the statement is extremely characteristic of you (very much like you) please select "Strong agreement". Of course, a statement may be neither extremely uncharacteristic nor extremely characteristic of you; if so, please select one of the choices in the middle of the scale that describes the best fit. Please respond as honestly as possible.

Table 41. Finalized RB-SOC Measure

Item		
1.	I read so I can have informed conversations with peers/co-workers.	
2.	I'm usually more concerned about finishing a text than comprehending it.	*
3.	If a text is difficult, I rarely re-read sections of the text.	*
4.	Reading is not my idea of fun.	*
5.	I read to learn more about a topic I'm unfamiliar with.	
6.	Understanding what I read is part of the fun of reading.	
7.	I'd rather read long texts than short ones.	
8.	When I read, I typically take note of key words or concepts in each paragraph.	
9.	I rarely read to relax.	*
10.	I like to read texts that don’t require much effort to figure out.	*
11.	One benefit of reading is that it can improve my vocabulary.	
12.	I often look up concepts I do not know.	
13.	I seldom make summaries of what I read.	*
14.	I read because I enjoy it.	
15.	I read so I can show others that I’m knowledgeable about a topic.	
16.	I often highlight or underline words to follow an author's line of thinking.	
17.	When I read, I nearly always want to understand what I've read.	
18.	I rarely make outlines to help me understand what I read.	*
19.	I do not feel that it is important to read regularly.	*
20.	It is important for me to read for fun, despite my other obligations.	
21.	If I really want to comprehend a book/article, I try to put myself in an environment that will help me do so (e.g., find a quiet place).	

22.	I try to avoid situations where there is a high chance I will have to read a text carefully.	*
23.	One reason I read is to improve my reasoning skills.	
24.	I'd rather read difficult texts than easy ones.	
25.	I'd rather spend effort understanding a long text than read a summary.	
26.	I often skim when I do not need to know the content very well.	*
27.	When I don't know words in a text, I try to figure out the meaning based on surrounding information.	
28.	I generally do not like reading dense texts.	*
29.	One reason I like to read is that it allows me to exercise my imagination and creativity.	
30.	I try to really understand a text, so I can form my own opinions about the content of the text.	
31.	I read to study for exams or a job promotion.	

Note. * denotes reverse scaling. Intrinsic reading goals, items 4, 6, 9, 14, 19, 120, 29; Extrinsic reading goals and learning strategies, items 8, 13, 15, 16, 18, 23, 31; Desire to understand, items 2, 3, 5, 11, 12, 17, 21, 27, 30; Desired reading difficulty, items 7, 10, 22, 24, 25, 26, 28. Likert scale ranged from 1 (Strong disagreement) to 7 (Strong agreement).

Appendix E

E1. Study 2 Assessment Distributions

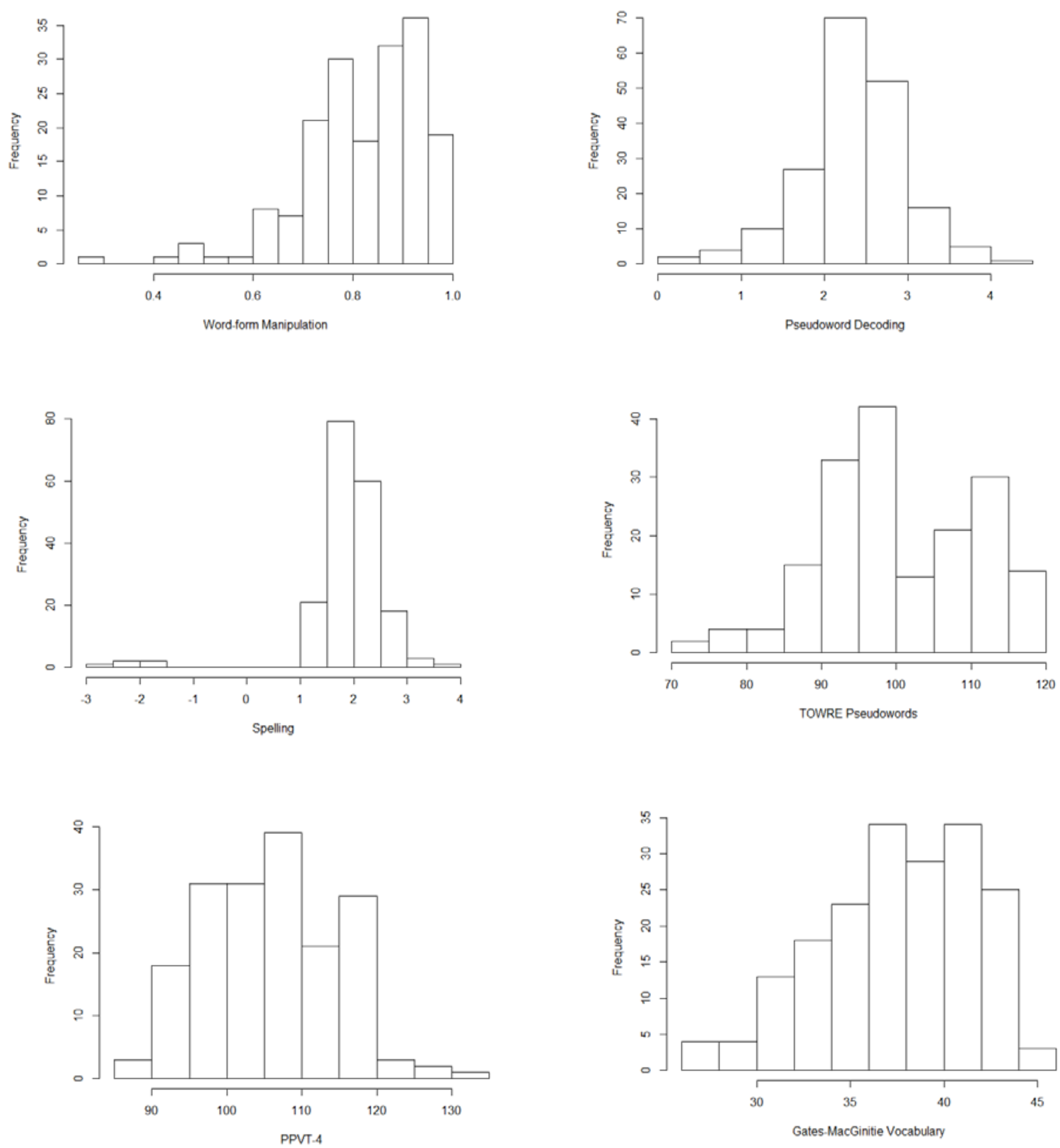

Figure 20. Histograms of Assessmensts in Study 2 (1)

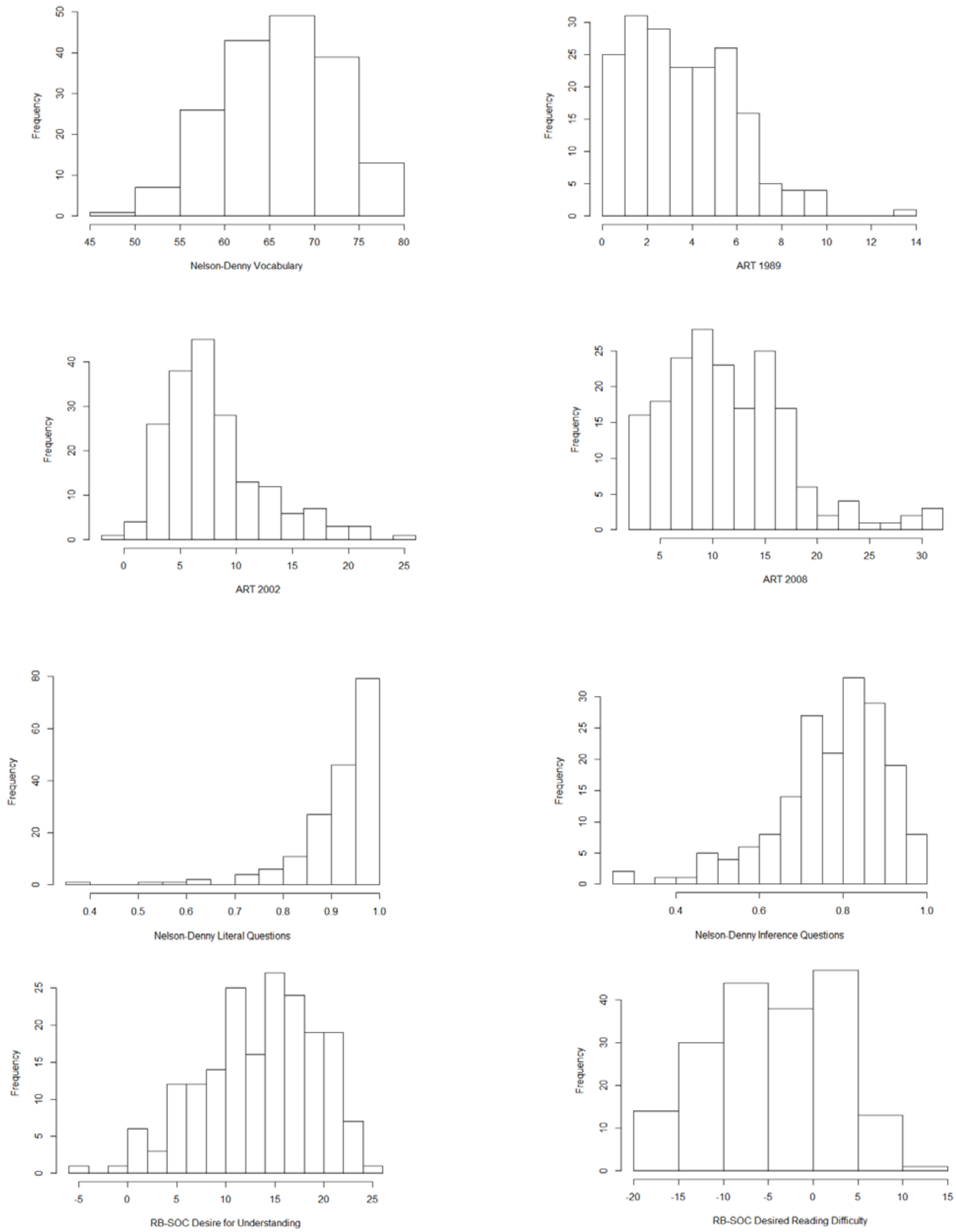

Figure 21. Histograms of Assessments in Study 2 (2)

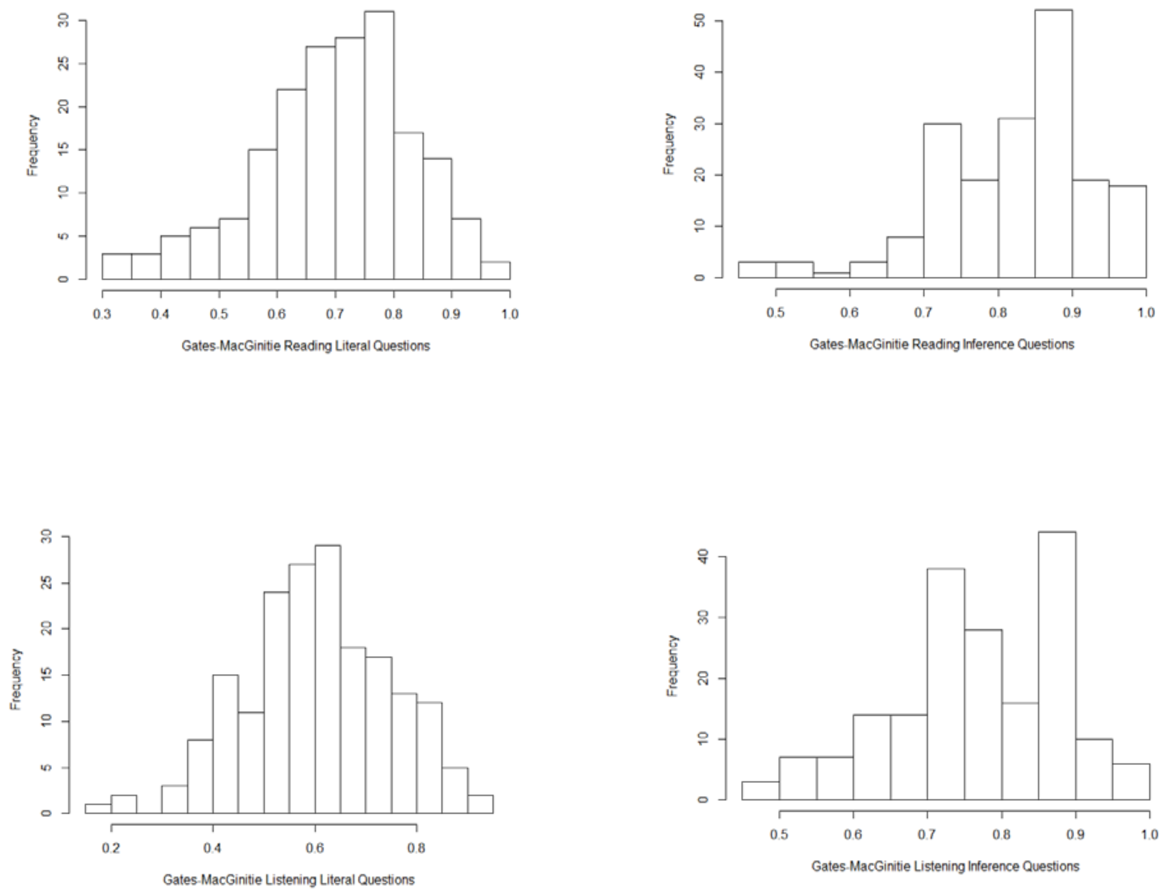

Figure 22. Histograms of Assessments in Study 2 (3)

Appendix F

F1. Newspaper Articles Rated in the Inference Norming Task

Table 42. List of Newspaper Articles in Study 3

Newspaper	**Title**	**Date**	**Topic**	**Sent.**	**Words**	**Reading level**
The Washington Times	A poverty-stricken community	January 13, 1895	Poverty	36	650	8.3
The Washington Times	The power of modern chemistry	November 4, 1897	Epidemic	37	653	10.2
The Washington Times	Infatuated man opens fire at a party	March 23, 1901	Murder/thriller	43	748	7.9
The Washington Times	A record Derby race	June 6, 1901	Sports	42	753	9
The New York Herald	Liberals to open fight for wine and beer at special session	November 17, 1922	Law & Treaties	41	731	11.7
The Washington Times	An epidemic	September 23, 1897	Epidemic	36	625	9.9
**The Washington Times*	Indignant over arrests	June 23, 1901	Law & Treaties	32	688	10.6
The Washington Times	Plan to end coal strike	September 9, 1902	coal strike; business deal	36	616	9.7
The New York Herald	Audit clerk steals $100,000 by posing as a shoe tycoon	August 20, 1921	Embezzlement & suing	36	789	11
The New York Herald	Owning a $115,000 Horse	June 29, 1921	Embezzlement & suing	40	600	7
The Washington Herald	World's champs capture opener from Nationals	April 12, 1912	Sports	45	672	6.5
The Washington Times	Death ends his struggle	March 27, 1901	Epidemic	50	814	7.8
The Washington Times	Caring for the homeless	May 5, 1901	Poverty	34	657	10.1
The Washington Times	A canal treaty	April 27, 1901	Law & Treaties	38	812	12

The Washington Times	Head of Illinois Wire Company slain	August 10, 1902	Murder/thriller	40	559	7.2
**The Washington Times*	Mile a minute in auto races	November 17, 1901	Sports	41	660	8
The Washington Times	A history of a State alcohol law	May 9, 1985	Law & Treaties	34	620	10
The Washington Times	Daly wins for the Army: Navy lose their annual football game with West Point Cadets	December 1, 1901	Sports	35	692	9
**The Washington Herald*	Ascends at 1 o'clock	June 5, 1907	Law & Treaties	38	607	8.6

Note. *Denotes the article contained one or more inferences that were rated as unlikely and were not included in in the final set of materials. Reading level is calculated as Flesch-Kincaid grade level.

Appendix G

G1. Interest End-of-study Questionnaire

1) During the task in which you answered questions about newspaper articles' use and our interest in them, did you prefer the listening or reading version?
    - Listening
    - Reading
    - Preferred both equally
2) Why do you think you preferred one over the other? Or if your preferred them equally, why?
3) Do you think it was easier to understand the articles when you were reading the newspaper articles or listening to them?
    - Listening
    - Reading
    - Understood both equally
4) If one was easier than the other, why do you think so?

Appendix H

H1. Example Passage

"World's Champs Capture Opener from Nationals" *The Washington Herald*. April 12, 1912.

Against Connie Mack's mighty Athletics, the baseball champions of the world, Manager Griffith's Washington Nationals opened the 1912 American League season (**control bridging: scored)**. When the runs were counted the Washington Nationals were on the wrong end of a 4 to 2 score (**bridging: lost)**.

The principal reason for the victory is very simple. Jack Coombs, the great "Iron Man" from Maine, allowed only four safe hits (**explicit middle: victory**). Not a member of Griffith's team reached third base until the seventh inning; also not a single safe hit was registered up to that time (**control paraphrase: attended**).

Walter Johnson, the Nationals' pitcher, was hardly at his best. Johnson himself cut loose with a ghastly wild heave in the fifth, which cost his team just two runs.

In justice to the Nationals, it must be said that the champs were handed all the breaks in luck (**explicit early: justice)**. This is usually the case when tailenders and champions clash on the ball field (**control bridging: ran**). Umpire Silk O'Loughlin was decidedly off color, and some of the strikes he muffed were costly (**explicit late: muffed**).

Twenty thousand people, the royalty of fandom, sat enwrapped in furs in the steel arms of that vast amphitheater (**paraphrase: stadium**). Mayor Blankenburg tossed out the first ball and the teams marched to the flagpole, headed by a brass band (**explicit early: tossed)**. In fact, everything possible was done to bid the Nationals welcome (**control predictive: sit)**. It followed in the natural course of events that the home team should win, and everybody in Quakertown is happy tonight with the exception of twenty-two bronzed athletes under the chaperonage of Clark Calvert Griffith (**explicit late: chaperonage)**.

A high wind and the bright sky bothered the outfielders on both teams (**paraphrase: sunny**). Jack Knight dropped a pop fly for the Nationals. Rube Oldring dropped two perfectly easy balls for the Athletics (**explicit late: easy**). Eddie Foster of the Nationals caught two vicious smashes and scored a run for the Nationals.

There was no question about the Athletics' first run in the opening round (**explicit late: round**). It was earned all right. Oldring made it home after Baker connected his only hit of the game to right field, his favorite spot (**control predictive: tackle**).

During the next three frames Walter Johnson, the Nationals' pitcher, simply toyed with the champions, but in the fifth inning the fortunes of war suddenly flopped (**explicit early: frames)**. Walter Johnson gathered up the ball and slammed to McBride at second to force Coombs to run. The ball is going yet. It sailed over McBride's head and past Milan, who watched the ball fall to the ground behind him (**bridging: missed).** Both Coombs and Strunk, who had been on first, scored on the error (**explicit late: error)**. This blunder just handed the champs the game. A perfect toss would have nailed Coombs at second. The Athletics' fourth run came in the sixth inning when Coombs hit a single.

In the seventh inning, Dan Moeller made Washington's first hit of the game (**explicit early: seventh)**. Flynn, who was up next to bat, crashed to the left field fence, over an outfielder's head and ran to second base. It was only by a lot of hustling by the outfielder that Moeller was held at third. Flynn played way off second to taunt the pitcher, and Moeller danced up and down the base lines between third and home. The pitcher, Coombs, thought he could get Flynn at the midway bag and cocked his arm (**predictive: throw).** However, Flynn slid back in safety, and Moeller dashed over the plate. Eddie Foster planted a pretty single over second and Flynn romped home. After that, the Athletics' team struck their third out **(control paraphrase: approval**). The score was 4 to 2 at the conclusion of the seventh inning.

From a Washington standpoint one feature was the fight and spirit shown by Griffith's charges (**explicit middle: fight)**. In the ninth inning everybody stood up and inwardly shivered. Could it be possible that the Nationals would really tie the score? There was Milan on third, Schaefer on second, and Jack Flynn, who previously scored for the Nationals, at the bat. Flynn drove the ball high in the air. Strunk was playing deep and had his glove in position (**predictive: catch)**. It was all over then but the shouting, and the crowd emitted a mighty roar of approval (**explicit middle: shouting)**.

Bibliography

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*(5), 1061–1070.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, *80*(3), 260–267.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295.

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289.

Baker, L. (1985). Differences in standards used by college students to evaluate their comprehension of expository prose. *Reading Research Quarterly*, *20*, 298-313.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(3), 336–345.

Bartholomew, D. J. (1980). Factor analysis for categorical data. *Journal of Royal Statistical Society*, *42*(3), 293-321.

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. *Journal of Educational Psychology*, *86*(2), 244–255.

Belloni, L. F., & Jongsma, E. A. (1978). The effects of interest on reading comprehension of low-achieving students. *Journal of Reading*, *22*(2), 106–109.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343–345.

Bernstein, M. R. (1955). Relationship between interest and reading comprehension. Journal of Educational Research, 49(4), 283–288.

Boersma, Paul & Weenink, David (2016). Praat: Doing phonetics by computer [Computer program]. Version 6.0.10, retrieved 8 January 2016 from http://www.praat.org/

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*(2), 122–159.

Boyles, N., & Scherer, M. (2012). Closing in on close reading. *On Developing Readers: Readings from Educational Leadership, EL Essentials*, 89-99.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., … Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of reading. *Reading and Writing*, *29*(3), 435–451.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243.

Brown, J., Fishco, V., & Hanna, G. (1993). *Nelson-Denny Reading Test: Manual for scoring and interpretation for forms G & H*. Chicago: Riverside.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306-307.

Calloway, R. C., & Perfetti, C. A. (2017). Integrative and predictive processes in text reading: the N400 across a sentence boundary. *Language, Cognition and Neuroscience*, *32*(8), 1001–1016.

Cain, K., Oakhill, J. V, Barnes, M. A, & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850–859.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and Component Skills. *Journal of Educational Psychology*, *96*(1), 31–42.

Carlson, S. E., van den Broek, P., McMaster, K., Rapp, D. N., Bohn-Gettler, C. M., Kendeou, P., & White, M. J. (2014). Effects of comprehension skill on inference generation during reading. *International Journal of Disability Development and Education*, *61*(3), 258–274.

Casteel, M. A. (1993). Effects of inference necessity and reading goal on children's inferential generation, *Developmental Psychology*, *29*(2), 346–357.

Chikalanga, I. (1992). A suggested taxonomy of inferences for the reading teacher. *Reading in a Foreign Language, 8*(2), 697-709.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, *51*(2), 291–294.

Coppens, K. M., Tellings, A., Schreuder, R., & Verhoeven, L. (2013). Developing a structural model of reading: The role of hearing status in reading development over time. *Journal of Deaf Studies and Deaf Education*, *18*(4), 489–512.

Coney, J. (2002). The effect of associative strength on priming in the cerebral hemispheres, *50*, 234–241.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311–325.

Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., & Kyle, K. (2017). Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes*, *54*(5), 340–359.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037//0012-1649.33.6.934

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277–299.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, *32*(3), 332–347.

Davis, M. H., Tonks, S. M., Hock, M., Wang, W., & Rodriguez, A. (2018). A review of reading motivation scales. *Reading Psychology*, *39*(2), 121–187.

Deci, E. L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 43-70). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Dunn, L. M., & Dunn, D. M. (1981). *Peabody picture vocabulary test-revised (PPVT-R). Forms L and M.* Circle Pines, Minn.: American Guidance Service.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, *79*(1), 262–300.

Elliot, J. A. (1999). Approach and avoidance motivation and achievement Goals. *Educational Psychologist*, *34*(3), 169–189.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). Intervention, evaluation, and policy studies: The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1–44.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.

Fincher-Kiefer, R. (1995). Relative inhibition following the encoding of bridging and predictive inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 981–995.

Finucci, J. M., Isaacs, S. D., Whitehouse, C. C., & Childs, B. (1982). Empirical validation of reading and spelling quotients. *Developmental Medicine and Child Neurology*, *24*(45), 733–744.

Finucci, J., Whitehouse, C., Isaacs, S., & Childs, B. (1984). Derivation and validation of a quantitative definition of specific reading disability for adults. *Developmental Medicine & Child Neurology*, *26*, 143–153.

Fisher, D., & Frey, N. (2012). Close reading in elementary schools. *Reading Teacher*, *66*(3), 179–188.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 465–495.

Fulmer, S. M., & Frijters, J. C. (2011). Motivation during an excessively challenging reading task: The buffering role of relative topic interest. *Journal of Experimental Education*, *79*(2), 185–208.

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository text on science and technology. *Scientific Studies of Reading*, *2*, 247–269.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., … Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, *96*(3), 403–423.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of the research. *Review of Educational Research*, *71*(2), 279–320.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., Carter, S.M., & Elliot, A.J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology*, *92*, 316-330.

Helder, A., Perfetti, C. A., van den Broek, P., Stafura, J. Z., & Calloway, R. C. (2019). ERP Indicators of local and global text influences on word-to-text integration. *Language, Cognition and Neuroscience*, *34*(1), 13–28.

Henk, W. A., & Melnick, S. A. (1995). The reader self-perception scale (rsps) – a new tool for measuring how children feel about themselves as readers. *Reading Teacher*, *48*(6), 470–482.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations, *Educational Psychology Review*, *13*(3), 191–209.

Horvath, M., Herleman, H. A., & Lee McKie, R. (2006). Goal orientation, task difficulty, and task interest: A multilevel analysis. *Motivation and Emotion*, *30*(2), 171–178.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767–787.

Juric, M. (2017). The role of the need for cognition in the university students' reading behaviour In *Proceedings of ISIC, the Information Behaviour Conference, Zadar, Croatia, 20-23 September, 2016: Part 2. Information Research*, *22*(1), paper isic1620).

Kamil, M.L., Borman, G.D., Dole, J., Kral, C.C., Salinger, T., and Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide.* NCEE No. 2008-4027. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, *23*, 115–126.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*(4), 765–778.

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*(1), 10–16.

Kim, Y. S. G. (2017). Why the simple view of reading is not simplistic: Unpacking component skills of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading*, *21*(4), 310–333.

Kim, Y. S. G., Vorstius, C., & Radach, R. (2018). Does online comprehension monitoring make a unique contribution to reading comprehension in beginning readers? Evidence from eye movements. *Scientific Studies of Reading*, *22*(5), 367–383.

Kingston, N., Wang, W., Davis, M. H., Tonks, S. M., Tiemann, G., & Hock, M. (2017). Development of the adaptive reading motivation measures: Technical report.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. (A. Kenny, David & T. D. Little, Eds.) (Third). New York: The Guilford Press.

Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *9*5(1), 22-40.

Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *NeuroImage*, *33*(1), 343–61.

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*, *23*(6), 701–717.

Lefly, D., & Pennington, B. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities*, *33*(3), 286–296.

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language*, *31*(5), 634–647.

Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. *Journal of Experimental Psychology*: *Learning Memory and Cognition*, *20*(6), 1456–1470.

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22–60.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests* (4th ed.). Itasca, IL: Riverside Publishing.

Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*(3), 273–284.

Mason, R. A., & Just, M. A. (2011). Differentiable cortical networks for inferences concerning people’s intentions versus physical causality. *Human Brain Mapping*, *32*(2), 313–329.

McDaniel, M. A., Waddill, P. J., Finstad, K., and Bourg, T. (2000). The effects of text-based interest on attention and recall. *Journal of Educational Psychology*, *92*(3): 492–502.

McKeown, M. G. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly*, *20*, 482–496.

Mckeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking Reading Comprehension. *Reading Research Quarterly*, *44*(3), 218–253.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *12*(1), 82–91.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1155–1172.

McKoon, G., & Ratcliff, R. (2015). Cognitive theories in discourse-processing research. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 42–67). Cambridge: Cambridge University Press.

McKoon, G., & Ratcliff, R. (2017). Adults with poor reading skills and the inferences they make during reading. *Scientific Studies of Reading*, *21*(4), 292–309.

McNamara, T. P. (2005). *Semantic priming: perspectives from memory and word recognition*. (H. L. Roediger, J. R. Pomerantz, A. D. Baddeley, V. Bruce, & J. Grainger, Eds.). New York: Psychology Press.

McWhaw, K., & Abrami, P. C. (2001). Student goal orientation and interest: Effects on students' use of self-regulated learning strategies. *Contemporary Educational Psychology*, *26*(3), 311–329.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, *53*(2), 253–279.

Murray, J. D., & Burke, K. A. (2003). Activation and encoding of predictive inferences: The role of reading skill. *Discourse Processes*, *35*(2), 81–102.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2–3), 131–157.

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology*, *91*(3), 488–496.

Nation, K. (2017). Nurturing a lexical legacy: reading experience is critical for the development of word reading skill. *Npj Science of Learning*, *2*(1), 3.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, *67*(3), 359–370.

National Research Council. (2012). *Improving adult literacy instruction: Options for practice and research*. National Academies Press.

Nell, V. (1988). The psychology of reading for pleasure: Needs and gratifications. *Reading Research Quarterly*, *23*(1), 6-50.

Nelson, M. J., & Denny, E. C. (1973). *The Nelson-Denny reading test*. Houghton Mifflin Company.

Ng, S., Payne, B. R., Steen, A. A., Stine-Morrow, E. A. L., & Federmeier, K. D. (2017). Use of contextual information and prediction by struggling adult readers: Evidence from reading times and event-related potentials. *Scientific Studies of Reading*, *21*(5), 359–375.

O'Brien, E. J., Rizzella, M. L., Halleran, J. G., & Albrecht, J. E. (1998). Updating a situation model: A memory-based text processing View. *Journal of Experimental Psychology*, *24*(5), 1200–1210.

Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–277). Baltimore: Paul H. Brookes Publishing Co.

Olson, R., Wise, B., Conners, F., Rack, J., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Quarterly Journal of Experimental Psychology*, *22*(6), 399–348.

Parault, S. J., & Williams, H. M. (2009). Reading motivation, reading amount, and text comprehension in deaf and hearing adults. *Journal of Deaf Studies and Deaf Education*, *15*(2), 120–135.

Partnership for Assessment of Readiness for College and Careers. (2011). *PARCC model content frameworks: English language arts/literacy grades 3–11*.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). New York, NY: Oxford University Press.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383.

Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition*, *7*(4), 273–282.

Perfetti, C. A., & Hart, L. A. (2001). The lexical basis of comprehension skill. In D. S. Gorfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti, C. A., Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading and comprehension skills. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–253). Oxford, UK: Blackwell.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37.

Perfetti, C. A., & Stafura, J. Z. (2015). Comprehending implicit meanings in text without making inferences. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 1–18). Cambridge: Cambridge University Press.

Perfetti, C. A., Yang, C.-L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*(3), 303–318.

Pintrich, P. R., (1989). The dynamic interplay of student motivation and cognition in the college classroom. In C. Ames & M. Maehr (Eds.), *Advances in motivation and achievement. Vol. 6: Motivation enhancing environments* (pp. 117–160). Greenwich, CT: JAI.

Potts, R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, *27*, 399–415.

Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology*, *80*(3), 319–323.

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2007). Development of lexical mediation in the relation between reading comprehension and word reading skills in Greek. *Scientific Studies of Reading*, *11*(3), 165–197.

Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The Components of the Simple View of Reading: A Confirmatory Factor Analysis. *Reading Psychology*, *33*(3), 217–240.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/

Ryan, R.M., Connell, J.P., & Plant, R.W. (1990). Emotions in non-directed text learning. Learning and Individual Differences, *2*, 1-17.

Sabatini, J. P., O'Reilly, T., Halderman, L. K., & Bruce, K. (2014). Integrating scenario-based and component reading skill measures to understand the reading behavior of struggling readers. *The Division for Learning Disabilities of the Council for Exceptional Children*, *29*(1), 36–43.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Schutte, N. S., & Malouff, J. M. (2007). Dimensions of reading motivation: Development of an adult reading motivation scale. *Reading Psychology*, *28*(5), 469–489.

Seifert, C. M., Robertson, S. P., & Black, J. B. (1985). Types of inferences generated during reading. *Journal of Memory and Language*, *24*(4), 405–422.

Shankweiler, D., Dreyer, L. G., Lundquist, E., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading and Writing: An Interdisciplinary Journal*, *8*, 267–294.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., … Bennett, A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading*, *3*(1), 69–94.

Shirey, L. L., & Reynolds, R. E. (1988). Effect of interest on attention and learning. *Journal of Educational Psycholog*y, *80*(2), 159–166.

Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Willson, V., … Byrns, G. (2010). A comparison of multiple-strategy methods: Effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. *Journal of Research on Educational Effectiveness*, *3*(2), 121–156.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior*, *22*(4), 437–448.

Singer, M., & Spear, J. (2015). Phantom recollection of bridging and elaborative inferences. *Discourse Processes*, *52*(5–6), 356–375.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *22*, 360–407.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure. *Memory & Cognition*, *20*(1), 51–68.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402–433.

Talwar, A., L. Tighe, E., & Greenberg, D. (2018). Augmenting the simple view of reading for struggling adult readers: A unique role for background knowledge. *Scientific Studies of Reading*, *22*(5), 351–366.

Taraban, R., Rynearson, K., & Kerr, M. (2000). College students' academic performance and self-reports of comprehension strategy use. *Reading Psychology*, *21*(4), 283–308.

Taylor, J. N., & Perfetti, C. A. (2016). Eye movements reveal readers' lexical quality and reading experience. *Reading and Writing*, *29*(6), 1069-1103.

Torgeson, JK., Wagner, RK., & Rashotte, C.A. (1999) Test of Word Reading Efficiency (TOWRE). Austin, TX: ProEd.

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, *45*(5), 453–466.

Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. (2005). The computational implementation of the landscape model: Modeling inferential processes and memory representations of text comprehension. *Behavior Research Methods*, *37*(2), 277–286.

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(3), 832–842.

van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text. In M. T. McCrudden, J. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 125–139). Greenwich, CT: Information Age Publishing.

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081–1087.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5–6), 360–372.

van den Broek, P., Risden, K., & Husebye-Hartman, E. (1995). The role of reader's standards of coherence in generation of inferences during reading. In E. P. Lorch & E. J. O'Brien (Eds.), *Sources of Coherence in Reading* (pp. 353–374). Mahwah, NJ: Erlbaum.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wagoner, S. A. (1983). Comprehension monitoring: What is it and what we know about it. *Reading Research Quarterly*, *18*(3), 328–346.

Wigfield, A. & Guthrie, J.T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, *89*(3), 420-432.

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, *45*(5), 432–445.

Williams, J. P. (2005). Instruction in Reading Comprehension for Primary-Grade Students: A Focus on Text Structure. *The Journal of Special Education*, *39*(1), 6–18.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension with children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, *50*, 1093–1109.

Yeari, M., van den Broek, P. & Oudega, M. (2015). Processing and memory of central versus peripheral information as a function of reading goals: Evidence from eye movements. *Reading and Writing*, *28*, 1071–1097.

Zabrucky, K., & Ratner, H. H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Reading Behavior*, *24*(3), 373–391.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 920–933.