Understanding How Language-Specific and Domain-General Resources Support Comprehension

Ryan D. Kopatich, Joseph P. Magliano, Keith K. Millis, Christopher P. Parker & Melissa Ray

Department of Psychology

Northern Illinois University

Accepted for publication in Discourse Processes:

Ryan D. Kopatich, Joseph P. Magliano, Keith K. Millis, Christopher P. Parker & Melissa Ray (2019) Understanding How Language-Specific and Domain-General Resources Support Comprehension, Discourse Processes, 56:7, 530-552, DOI: 10.1080/0163853X.2018.1519358

Discourse Processes peer review process: Manuscripts submitted to this journal undergo editorial screening and peer review by anonymous reviewers.

Acknowledgments:

This work was supported by the Institute of Educational Sciences (under Grants R305A150193 and R305G040055) and the Center for the Interdisciplinary Study of Language and Literacy (CISLL) at Northern Illinois University.

Abstract

A large body of work has demonstrated that reader resources influence inference processes and comprehension, but few models of comprehension have accounted for such resources. The Direct and Mediational Inference model of comprehension (DIME) assumes that general inference processes mediate the effects of reader resources on general comprehension proficiency.

The current study proposes an extension of DIME, the Inferential Mediation Model (IMM), to account for comprehension as it occurs while reading a particular text. College students were administered a battery of reader resource measures. In addition, they completed a think-aloud tool that measured comprehension and readers' inference processes while reading specific texts. A path analysis revealed that inference processes partially mediate the relationships between reader resources and comprehension performance. These results support the DIME and IMM, and suggest that inference processes that support mental model construction mediate the impact of reader resources on comprehension.

Understanding How Language-Specific and Domain-General Resources Support Comprehension

Readers construct a mental model of texts, which is the basis of comprehension (Kintsch, 1998; 1988). A coherent mental model is a representation of what is depicted and implied by the text in which ideas stated in the text are semantically connected (McNamara & Magliano, 2009b). One important process in establishing coherence is the generation of inferences that establish how discourse constituents are connected as well as how relevant background knowledge is activated and incorporated into the model (e.g., Graesser, Singer, & Trabasso, 1994). Moreover, when building a mental model, readers may rely on a variety of cognitive resources. These resources can be specific to text comprehension, such as basic reading skills and vocabulary (Farley & Elmore, 1992; Nagy, 2007) or can be relatively general, such as working memory, metacognitive awareness, and need for cognition (Dai & Wang, 2007; Just & Carpenter, 1992). Although there is an extensive literature showing that general and specific resources influence comprehension (Acheson, Wells, & MacDonald, 2008; MacDonald, Just, & Carpenter, 1992) and inference generation (Calvo, 2005; Singer, Andrusiak, Reisdorf, & Black, 1992), more research is needed to explore and test potential relations among reader resources and inference processes in terms of creating mental models, and thus supporting comprehension.

The goal of the present study was to explore the relation between reader resources and inference generation in terms of how their relationship with comprehension. Our theoretical framework is aligned with the Direct and Mediational Inference Model of reading comprehension (DIME; Ahmed et al., 2016; Cromley & Azevedo, 2007). DIME assumes that reader resources and inference ability have direct effects on comprehension and that reader resources also have indirect effects through inference ability. Below we discuss our framework in the context of DIME.

The DIME and IMM Models

The DIME model assumes that a variety of reader resources (i.e., background knowledge, reading strategies, word reading, and vocabulary) support comprehension directly, and also indirectly support comprehension via inferences (Cromley & Azevedo, 2007). To test their model, Cromley and Azevedo (2007) administered a series of assessments to 9th grade students that were aligned with language-specific (e.g., vocabulary, word reading) and domain-general resources (e.g., metacognitive skills, prior knowledge), general inference skill, and comprehension proficiency. They found that a partially mediated model provided the best fit of the comprehension proficiency data, in which the impact of language-specific and domain-general resources on comprehension proficiency was partially mediated through general inference ability. Ahmed and colleagues (2016) also tested the DIME model in a sample of students ranging from 7th to 12th grade, and similarly found that a partially mediated model best accounted for performance on assessments of general comprehension proficiency. Finally, Cromley, Snyder-Hogan, and Luciw-Dubas (2010) also found evidence for the DIME model in the context of science texts when testing undergraduate science majors.

All of the research on DIME to date has examined general inference skill and comprehension proficiency in a decontextualized setting. What we mean by “decontextualized setting” is that the inference tasks were not directly related to the comprehension outcome measures. That is, instead of inferences about the passages in which comprehension was measured, a multiple-choice measure was used where participants read brief passages and were asked directly about which option reflected a correct inference within the text. Their test of inference ability was intended to reflect participants’ inference skill that should generalize to other texts. However, in the present study, we were interested in testing a model inspired by the

DIME in the context of inference processes generated while reading specific texts and determining the extent to which these processes mediate the impact of reader resources on comprehension performance for those texts. Our perspective is that reader resources are related to the inferences that are generated to support mental model construction and comprehension as texts are processed, and relying on decontextualized, general measures of inference ability does not afford exploring this possibility.

In the present study, we rely on verbal protocols (i.e., think-aloud protocols) produced during reading to provide an assessment of the inferences that were generated for specific texts (e.g., Linderholm & van den Broek, 2002; McMaster et al., 2012; Trabasso & Magliano, 1996; van den Broek, Lorch, Linderholm, & Gustafson, 2001). Additionally, we relied on open-ended, how and why comprehension questions to provide an assessment of how well those texts were comprehended, which provided the outcome measure. As will be described below, answering the questions required readers to establish causal connections between text ideas that were important to establish coherence, and therefore, this measure should be sensitive to the quality of a mental model because these relationships should be stored in them (Graesser & Franklin, 1990; Magliano et al., 2011).

We propose a variant of the DIME which we call the Inferential Mediation Model (IMM (See Figure 1 for the IMM; Ahmed et al., 2016; Cromley & Azevedo, 2007). Mental models are incrementally updated and thus comprehension is incrementally established as a text is read (Gernsbacher, 1990; Kintsch, 1998; 1988; Zwaan & Radvansky, 1998). Assuming they are accurate, the inferences generated as one reads through a text are a crucial component of the comprehension process and support the construction of a mental model. This highlights a key difference between previous research on DIME and the IMM. The DIME model was proposed to explain the relationships between reader resources, general inference proficiency, and general comprehension proficiency (as measured by standardized reading comprehension tests, such as the Gates-MacGinitie; Ahmed et al., 2016; Cromley & Azevedo, 2007).

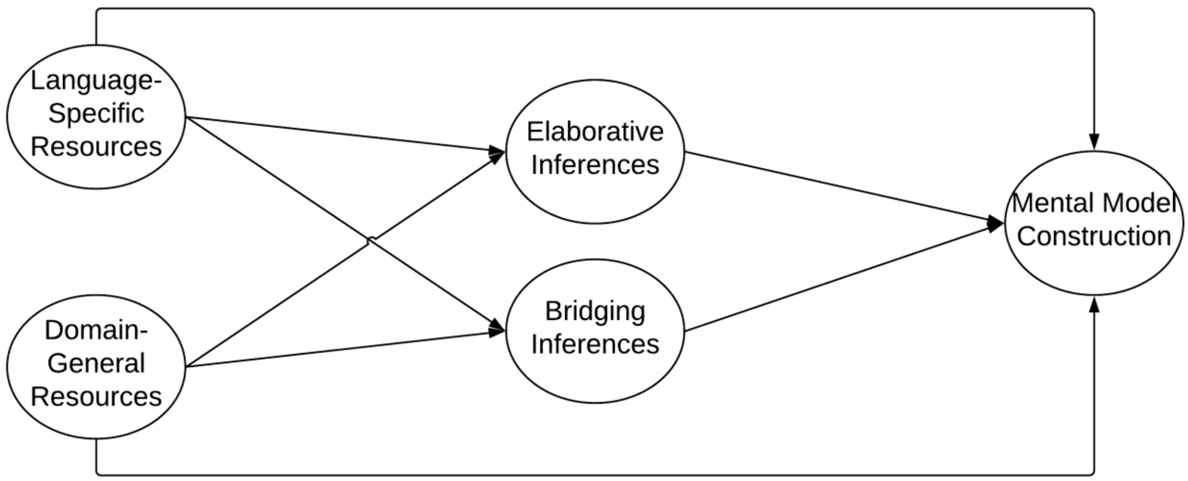

Figure 1. Proposed IMM tested by the current study. The IMM proposes that the effect of domain-general and language-specific resources on mental model construction is partially mediated by inference generation processes.

IMM extends DIME by focusing on the relationships between readers' resources, the inferences generated while reading specific texts, and the mental model that is constructed which results in comprehension.

We focus on two broad classes of inferences that have been previously found to be related to mental model construction: bridging inferences and elaborative inferences (Magliano & Millis, 2003; Magliano, Millis, The RSAT Development Team, Levinstein, & Boontum, 2011; Millis, Magliano, & Todaro, 2006; Singer et al., 1992; Whitney, Ritchie, & Clark, 1991). Bridging inferences establish how discourse components are connected, whereas elaborative inferences establish how background knowledge can be incorporated into the mental model (McNamara & Magliano, 2009b; Singer, 1988). This process can be seen in Figure 1. Mental models are constructed through a set of iterative processes that support inference generation and operate as each clause and sentence is read (e.g., Kintsch, 1998; 1988). Comprehension thus emerges in the sense that the inference processes contribute to how the mental model of the text is created and updated as the reader progresses through the text (Zwaan & Radvansky, 1998).

Language-Specific Resources

For the current study, we define language-specific resources to be the skills and experiences that are directly tied to reading or are linguistic in nature. There is a plethora of skills that have been explored in the reading literature, including fluency (e.g., Klauda & Guthrie, 2008), word reading ability (e.g., Ouellette, 2006; Perfetti & Hogaboam, 1975), inference ability (Danemen & Hannon, 2001; Magliano & Millis, 2003; Magliano et al., 2011), and exposure to text (Cunningham & Stanovich, 1991). Our choice of measures was constrained by three considerations: consistency with the DIME model, relevance for post-secondary readers, and time constraints regarding testing. Relevance of constructs was a crucial consideration because some measures of language-specific resources, such as reading fluency and phonemic awareness, predict comprehension in younger readers (Ehri et al., 2001; National Center for Education Statistics, 2015), but have not been found to be predictive in post-secondary readers. The length of testing time has obvious implications on the number and length of the assessments we could administer. As such, we wanted to use measures that tapped into constructs relevant to college readers but that could be administered in a timely manner.

Because of these constraints, the constructs of vocabulary knowledge, general comprehension proficiency, and exposure to text were included as indicators of language-specific resources in the present study. Vocabulary knowledge is an explicit language-specific component of the DIME model, which predicted significant variance in their measures of inference ability and comprehension proficiency (Ahmed et al., 2016; Cromley & Azevedo, 2007). Cromley and Azevedo (2007) based their argument of the importance of vocabulary on a rich literature showing that interventions that target vocabulary typically lead to robust improvements in reading comprehension (Beck, McKeown, & Kucan, 2013; Beck, Perfetti, & McKeown, 1982; Calvo, 2005; Medo & Ryder, 1993; Perfetti, McKeown, & Kucan, 2010; Perfetti & Stafura, 2014). There are many reasons why vocabulary training improves comprehension, but one obvious reason is that a relatively rich vocabulary supports the activation of linguistic knowledge (i.e., activation of concepts in the lexicon; Beck et al., 1982; Perfetti et al., 2010; Perfetti & Stafura, 2014) and general world knowledge that supports the construction of a coherent mental model (Kintsch, 1998).

In addition to vocabulary, another candidate for a language-specific resource is general comprehension proficiency, as measured by standardized tests of comprehension such as the Gates-MacGinitie reading test (MacGinitie, 2000). General comprehension proficiency is an important predictor of mental model construction (e.g., Long & Chong, 2001; Magliano & Millis, 2003). Given that our interest is in how readers construct a mental model for a given text, rather than comprehension as a general skill, we include this as a predictor of the quality of the mental model built for a particular text. In line with previous literature, we predict that this construct will help to account for individual differences in how well people generate inferences and the quality of mental models people generate for a specific text (e.g., Magliano & Millis, 2003).

As might be expected, the more one reads, the better their vocabulary will be and the more proficiently they will read (Cunningham & Stanovich, 1991; McBride-Chang, Manis, Seidenberg, Custodio, & Doi, 1993; Stanovich & Cunningham, 1993). As such, in order to fully account for language-specific skills, one must consider factors such as the exposure one has to various texts. Furthermore, this measure could also be conceived as being related to the knowledge construct that has been used in the DIME model (Cromley & Azevedo, 2007).

Domain-General Resources

Broadly speaking, domain-general resources are cognitively complex processes that can be deployed to assist the comprehension of language, but are not bound to language. As mentioned before, we are interested in constructs that are aligned with DIME, contribute to the construction of mental models in a post-secondary sample, and are practical given the constraints of the design of the study. For the current study we were interested in the contributions of working memory, metacognitive strategies, and need for cognition on inference generation and mental model construction.

Working memory capacity has been shown to be a consistent predictor of comprehension (Calvo, 2005; Daneman & Carpenter, 1980; Fedorenko, Gibson, & Rohde, 2006; Just & Carpenter, 1992; MacDonald & Christiansen, 2002; MacDonald et al., 1992). Because working memory is associated with the activation, suppression, and maintenance of information (Engle, 2002), it is little surprise that it has often been linked to reading comprehension (e.g., Daneman & Carpenter, 1980; Just & Carpenter, 1992). Together, these processes (activation, suppression, maintenance) allow readers to update information in mental models and to shift from one to another (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, 2002; Gernsbacher, 1990; McVay & Kane, 2012; Oberauer, Süß, Wilhelm, & Wittman, 2003). The research on working memory and comprehension suggests that in order to understand how mental models are constructed and updated, working memory's contribution to inference generation and comprehension within a given text must be studied.

Updating information in mental models also requires some degree of awareness of when a mental model needs to be updated and the strategies needed to update the mental model. Awareness of the process is especially acute when comprehension fails. Thus, metacognition is another crucial domain-general resource in the process of mental model construction. Metacognition refers to an individual's awareness and deliberate conscious control over his or her cognitive activities (Hacker, 1998). It is a dynamic process in which the goals of learning maintain selective processing, provide an ongoing evaluation of progress towards these goals, and drive engagement in selective updating to ensure learning is on track (McNamara & Magliano, 2009a; Thiede, Griffin, Wiley, & Redford, 2014). Metacognition has two important dimensions. First, *metacognitive awareness* is essential to students' response to a task, and it includes knowledge of one's cognitive processes and outcomes, as well as knowledge of the strategies needed to accomplish that task (e.g., error detection and repair; Paris, Wasik, & Turner, 1991). Second, *metacognitive knowledge and behaviors* enable readers to flexibly employ a repertoire of known comprehension strategies, suited to a variety of academic reading tasks, to enable their understanding of texts (Hacker, 1998; Mokhtari & Reichard, 2002; Schmitt, 1988). To date, all measures of metacognition are grounded in activities such as reading (e.g., Mokhtari & Reichard, 2002), which may make measures of metacognition less domain-general. However, the construct of metacognition is clearly domain-general as it can be applied to processes that occur outside of reading (e.g., self-regulation; Fox & Riconscente, 2008).

Finally, while many aspects of the comprehension process may be tacit (e.g., Kintsch, 1988; McKoon & Ratcliff, 1992; Myers & O'Brien, 1998), reading is an inherently deliberate, and goal-directed activity (e.g., Graesser et al., 1994). Even when readers make a cursory glance at text, they are deciding to engage in some form of processing and they decide what level of understanding they wish to achieve (e.g., van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011; van den Broek et al., 2001). This is particularly the case in the context of academic reading (Guthrie, McGough, Bennett, & Rice, 1996; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). Successful readers, and in particular readers who are successful in academic contexts, are intrinsically motivated to learn, and typically demonstrate a high need for cognition (Dai & Wang, 2007; Guthrie et al., 1996). To specify, need for cognition refers to one's propensity to engage in cognitively demanding tasks (i.e., how often they do it, how much they like to, etc.). Thus, a person's need for cognition is also an important general resource that a reader brings to a reading situation.

It should be noted that these domain-general resources are often used in studies of linguistic phenomena (e.g., Acheson & MacDonald, 2009; MacDonald & Christiansen, 2002). However, for the current study, we argue that these constructs are general in the sense that, while they may be related to language comprehension (e.g., Daneman & Carpenter, 1980), they are related to meaning making processes in non-linguistic domains (e.g., visual information processing; Colzato, Spapé, Pannebakker, & Hommel, 2007). Moreover, in the present study we tried to avoid domain-general instruments that are directly linked to language processing. For example, we used the Operational Span task (OSPAN) to assess working memory capacity, which is intended to be a relatively domain-general assessment of working memory span. However, most instruments of metacognition are grounded in specific activities, and therefore it is not possible to have a truly domain-general assessment of this construct. Nonetheless, a principal components analysis allowed us to examine whether that instrument loaded on the same component associated with domain-general resources more robustly than a component associated with language-specific resources.

Reader Resources and Inference Generation

There is an extensive literature on inference generation, the processes that support inferences, and their role in mental model construction (Graesser, et al., 1994; McKoon & Ratcliff, 1986; 1992). In fact, theories of discourse comprehension universally assume that inferences are essential for establishing coherence during reading (McNamara & Magliano, 2009b). However, in this review, we primarily focus on the role of reader resources in the construction of inferences. What evidence is there that language-specific and domain-general resources are related to bridging and elaborative inferences? There is much research that suggests that the language-specific construct of general reading proficiency (Cain, Oakhill, Barnes, & Bryant, 2001; Long, Oppy, & Seely, 1997; Long, Seely, Oppy, & Golding, 1996; McNamara, de Vega, & O'Reilly, 2007; Magliano & Millis, 2003) and the domain-general resource of working memory (Budd, Whitney, & Turley, 1995; Calvo, 2005; Daneman & Carpenter, 1980; Just & Carpenter, 1992; Linderholm, 2002; Virtue, Parrish, & Jung-Beeman, 2008; Whitney et al., 1991) are positively correlated with bridging and elaborative inferences. However, the relationship between inferences and domain-general resources may be complex.

While individual differences in working memory resources are correlated with inference generation (e.g., Calvo, 2005; 2001; Whitney et al., 1991), the nature of the relationship may be different for bridging and elaborative inferences (Fincher-Kiefer, 2001; Fincher-Kiefer & D'Agostino, 2004; Whitney et al., 1991). Fincher-Kiefer and D'Agostino (2004) had participants read texts that required a bridging or predictive inference (which is a type of elaborative inference) under a working memory load. Across several experiments, they found that the working memory load disrupted the generation of predictive inferences, but not bridging inferences. They argued that predictive (elaborative) inferences in general require effortful processing because they are generated relatively rarely. On the other hand, bridging inferences were highly practiced because they are necessary to establish coherence, and therefore, was less affected by the load.

However, Magliano, Larson, Higgs, and Loschky (2016) found that a working memory load disrupted bridging inferences in the context of visual narrative. They argued that the discrepancy in their findings with Fincher-Kiefer and D'Agostino (2004) may have stemmed from the load task. Magliano et al. (2016) used a load task that required remembering the order of the presentation of the information associated with that task. In contrast, Fincher-Kiefer and D'Agostino (2004) manipulated working memory load by having participants remember a set of individual letters. As such, the load in Magliano et al. (2016) may have been more disruptive of bridging inferences than tasks that do not require remembering order because bridging inferences, especially those for visual narratives, require making connections between sequenced discourse constituents.

Whitney et al., (1991) presented a study that suggested that high and low-(working memory) span readers may differentially produce bridging and elaborative inferences. They measured working memory capacity and had participants think-aloud while reading texts in order to identify the extent to which participants generated bridging and elaborative inferences in their protocols. They found that high-span readers tended to produce more bridging inferences than low-span readers, whereas low-span readers tended to produce more elaborative inferences than high-span readers. They argued that high-span readers could access the prior text to support bridging inferences when thinking aloud because they had an enriched representation, whereas low-span readers relied on what general knowledge became activated after reading the sentence that preceded the think-aloud prompt.

Overview of Current Study and Research Questions

In order to test the IMM depicted in Figure 1, participants were asked to think-aloud while reading 6 short texts. In addition, participants were given measures associated with language-specific (measures of vocabulary, comprehension proficiency, text exposure) and domain-general (working memory, metacognitive awareness, and need for cognition) resources. Think-aloud procedures have previously been used to assess the extent to which participants engage in bridging and elaborative inference generation in specific texts (Coté & Goldman, 1999; Magliano & Millis, 2003 Magliano et al., 2011; McMaster et al., 2012; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007; Trabasso & Magliano, 1996; van den Broek, 1994). Moreover, inferences revealed when thinking aloud are correlated with performance on a variety of measures of comprehension (Magliano & Millis, 2003; Millis et al., 2006; Trabasso & Suh, 1993; van den Broek, 1994; van den Broek et al., 2001). We assessed comprehension with adjunct, open-ended why and how questions because answers to these kinds of questions have been shown to be sensitive to comprehension level (Graesser & Clark, 1985; Graesser & Franklin, 1990)

The current study was motivated by three research questions. First, do measures that are domain-general or language-specific share common variance with each other? Based on previous literature, we expected that there would be common variance shared among measures of basic comprehension skill, vocabulary knowledge, and text exposure (language-specific resources) and among measures of working memory, metacognitive awareness, and need for cognition (domain-general resources). A principal components analysis (PCA) was used to derive latent variables that corresponded to language-specific and domain-general resources, which were then used in analyses to answer the two subsequent research questions.

Second, does inference generation account for any variance in comprehension above and beyond that explained by language-specific and domain-general resources? There is already some evidence that inference generation predicts comprehension beyond language-specific resources. Specifically, Magliano and Millis (2003) found that inference generation accounts for more variance in comprehension than standardized tests of reading. However, the current study involves a more comprehensive examination of the kinds of resources that affect readers' comprehension of a text.

Finally, does the data support the partially mediated IMM model for contextualized inferences? We expect that this model will fit the data well based on previous evidence that links resources such as working memory, vocabulary, and general reading proficiency to inference generation (Calvo, 2005; Cromley & Azevedo, 2007; Cutting & Scarborough, 1991; Farley & Elmore, 1992; Just & Carpenter, 1992; Magliano & Millis, 2003)—as well as the research linking inferences to comprehension (Graesser et al., 1994; McNamara & Magliano, 2009b; Zwaan & Radvansky, 1998).

Method

Participants

One hundred forty-eight students from Northern Illinois University (NIU) participated in this study in exchange for course credit in an introductory psychology course. The current study utilized archival data, in which no racial or gender demographic information was recorded, so these can only be extrapolated from institutional data. The introductory psychology course fulfills a general education requirement, and thus, is likely representative of NIU's student demographics. NIU is a diverse university with a majority white population (57%) and significant black (16%), Hispanic (15%), and Asian (5%) populations and the student body has roughly equal male and female segments (49% female).[1]

Measures and Materials

Language-specific measures. There were three assessments of language-specific resources. The Gates-MacGinitie reading test (Fourth Edition, Level 10/12, Form T) served as an indicator of general comprehension proficiency. This form of the Gates-MacGinitie reading test has high reliability (K-R 20's > .88; test-retest r's > .58) and is highly correlated with other measures of reading ability (MacGinitie, 2000).

The frequency that an individual engages in reading and the breadth of their reading experience was measured using the Author Recognition Test[2] (Stanovich & West, 1989). The Author Recognition Test is a measure designed to assess exposure to text. The test consists of a list of 80 names in checklist form. Forty of the names are actual authors and 40 are "foils," which are names of people who are not authors. Participants are asked to place a check next to the names of authors whom they recognize. Scoring is done by taking the number of items correctly marked and subtracting the number of foils. Possible scores range from -40 to 40. Reliability is evidenced by high Cronbach's alpha (.92; Cunningham & Stanovich, 1997) and good split-half reliability (.86; Stanovich & Cunningham, 1993).

Finally, vocabulary knowledge was assessed with the Shipley Institute of Living Scale (Zachary & Shipley, 1986). The vocabulary subtest was used as a measure of vocabulary skill.

---

[1] Unfortunately, demographic information about the sample used in this study was not available. The data were collected over a decade ago, when it was not common practice to report detailed demographic data. As such, we provided information about the population of NIU from which the sample was derived.

[2] Although there is an updated version of the Author Recognition Task (i.e., Acheson, Wells, & MacDonald, 2008), the data used in the current study were collected before this updated version was published.

This subtest consists of 40 multiple-choice questions in which the respondent is asked to choose which of four words is closest in meaning to a target word. Participants were allowed to work at their own pace. Raw scores range from 0 to 40. Test-retest reliability has been found to be quite high for this measure (r = .89; Nixon, Parsons, Schaeffer, & Hale, 1995), and it has been used extensively in the literature as a measure of vocabulary ability (e.g., Darowski, Helder, Zacks, Hasher, & Hambrick, 2008; Vasquez, Binns, & Anderson, 2017; Was, Rawson, Bailey, & Dunlosky, 2011).

Domain-general measures. There were three measures of domain-general resources. The Need for Cognition Scale (Cacioppo & Petty, 1982) was used to assess the participants' tendency to engage in and enjoy thinking. This measure consists of 18 items (e.g., “I find satisfaction in deliberating hard and for long hours”) in which participants are asked to rate themselves on a 9 point Likert-type scale ranging from -4 (very strong disagreement) to +4 (very strong agreement). Higher scores indicate a greater need for cognition. Possible scores range from -72 to 72. The reliability of the measure has been studied extensively (for a review, see Cacioppo, Petty, Feinstein, & Jarvis, 1996) and Cronbach’s alpha has been consistently high for the measure (.62 < α’s < .92; Cacioppo et al., 1996).

To test working memory span, participants completed the operation span (OSPAN) task (Turner & Engle, 1989; Unsworth, Heitz, Schrock, & Engle, 2005). In this task, the participant is presented with sets of multiple trials in which a mathematical operation is presented and followed by a random letter of the alphabet (e.g., IS 9 / 3 + 1 = 6? K). The participant is required to decide if the mathematical operation is true or false, indicate their decision by clicking the corresponding option on the screen, and to remember the letter for later recall. Blocks consisted of two to five trials and a practice session was administered before the test trials. An individual’s working memory capacity was computed as the number of letters the participant could recall in the order that they were presented by selecting them from a matrix of possible letters. Possible scores range from 0 to 75. Internal consistency for this measure has been shown to be acceptable (α = .70-.80; Conway et al., 2002; Unsworth et al., 2005).

Finally, in order to assess knowledge and use of metacognitive strategies employed by the reader, the Metacomprehension Strategy Index was administered (MSI; Schmitt, 1988). The MSI is a 25-item multiple-choice questionnaire that assesses an individual's awareness of his or her reading process as well as his or her awareness of reading strategies. Participants are given a scenario ("Before I begin reading, it's a good idea to…") and are asked to select the best strategy of the choices given (e.g., "make some guesses about what I think will happen in the story"). Possible scores range from 0 to 25. It has been shown to have good reliability (KR-20 = .87; Schmitt, 1990; α = .82; O'Reilly & McNamara, 2007) and is strongly correlated with other measures of metacomprehension (Schmitt, 1990). Additionally, it has been used in prior work on discourse comprehension as measure of metacognitive strategies (McNamara, O'Reilly, Best, & Ozuru, 2006; O'Reilly & McNamara, 2007). The MSI is specific to text, and could arguably be considered a language-specific assessment. As mentioned earlier, to our knowledge most assessments of metacognition are grounded in some activity, such as reading. However, if this is a domain-general resource as argued above, the principal components analysis should show that it is more correlated with the latent variable associated with domain-general resources than the latent variable associated with language-specific resources.

Measures of inference generation and comprehension. The Reading Strategy and Assessment tool (RSAT; Magliano et al., 2011) was used to assess both comprehension and inference processing within specific texts. RSAT is a computer-based tool which involves having participants read a text silently one sentence at a time. After pre-selected target sentences (which are not known to the participants beforehand), participants are asked one of two different types of questions. Indirect questions are intended to elicit a think-aloud type responses ("What are you thinking now?"). Participants are instructed (and trained) to report any thoughts they have regarding their understanding of the sentence that they just read in the context of the text. Responses to indirect questions reveal bridging and elaborative inferences (Gilliam et al, 2007; Magliano et al., 2011). Direct questions are intended to provide an assessment of how well the test taker comprehends the text up to that point. As such, they were the criterion variable in the analyses presented below. These direct questions take the form of "why" and "how" questions specific to the sentence that occurred prior to the prompt, but which require readers to draw on prior discourse (e.g., the participant answers the question "How did the tumor develop?" after reading the sentence "A tumor develops."). The questions were designed such that the answers were contained in the prior text, and required the tests takers to have inferred causal relationships between the relevant discourse constituents. Only one sentence is presented at a time and no sentences were present while participants answered the indirect and direct questions. Thus, responding to indirect and direct question prompts requires access to the mental model for the text. Participants type their answers into a box on the computer screen that appears below the question prompt.

The texts used in this study constituted a form of RSAT that was developed by Magliano et al. (2011). This form contains six texts: two texts from the genres of fictional narratives, historical texts, and science texts each. The Flesch-Kincaid grade levels for the texts varied from 3.2 to 12.0, with the narrative texts having lower grade levels. Because of the difference in grade levels between genres, we did not analyze differences between genres as they were confounded with text difficulty. The texts varied in length from 22 to 39 sentences and from 343 words to 491 words. In all, there were 18 direct and 15 indirect questions across all six texts. Texts were presented in a random order for each participant. See Appendix for an example text. RSAT does have automated scoring protocols and there is validity and reliability data for those scoring systems (Magliano et al., 2011; Millis & Magliano, 2012). However, in the present study, these scoring systems were not used for the direct or indirect protocols. Rather, they were hand coded and as such, reliability was based on inter-rater reliability in executing the coding systems that were used to analyze the protocols.

Analysis of Protocols

The indirect answers were scored by trained human judges using a coding system that was developed by Magliano et al. (2011) and provided criteria for determining the presence of at least one bridging or elaborative inference. The unit of analysis was the entire answer to a question. Bridges were instances where people mentioned concepts and clauses from the prior text. Local bridges occurred when the answer contained information from the immediately prior sentence and distal bridges contained information from all other prior sentences. The coding system makes a distinction between bridging to verb predicates and arguments (see also McNamara, 2004). Specifically bridging to verb predicates was assumed to reflect that the reader was connecting the current sentence with an event or state described in a prior sentence(s), whereas bridging to arguments reflected making connections to entities and concepts discussed in the prior discourse context (Zwaan, Langston, & Graesser, 1995; Kurby, Britt, & Magliano, 2005). Both local and distal bridges were scored via the same criteria. A “0” indicated that the answer did not contain a local/distal bridge. A “1” indicated that the answer contained a noun or noun phrase from a prior sentence. A “2” indicated that the answer contained a verb clause from a prior sentence. Judges were trained to detect synonymous expressions for local and distal bridges. These scores were then summed to give an overall bridging score. Elaborations were instances of inferences that contained concepts not mentioned in the text and therefore were generated from their world knowledge. A “0” indicated that no elaboration was present. A “1” indicated that the answer contained a noun or noun phrase not present in the text. A “2” indicated that the answer contained a main idea from the text with a verb clause that was not stated by the text or was likely the result of reasoning (see Appendix for examples). Personal recollections (e.g., “There was a thunderstorm last night.”, “My grandma died of cancer”) and evaluative statements (e.g., “I hate thunderstorms.”, “Cancer is scary.”) were not considered elaborations which contribute to the mental model of the discourse. This decision was based on findings from Todaro, Magliano, Millis, McNamara, & Kurby (2008) who found that comprehension outcomes were negatively correlated with the production of recollections and evaluations when thinking aloud. Trained judges worked in pairs and there were two groups of trained judges. Inter-rater reliability for assessing the presence of each category of processing was acceptable (Cohen’s κ’s ranged from .80 to .93). The appendix contains samples of coded protocols. It is important to note that the processes are not mutually exclusive and any given protocol could contain both bridging and elaborative inferences (e.g., McNamara, 2004; Trabasso & Magliano, 1996). Because of the coding system used (i.e., looking at noun and verb phrases, irrelevant elaborations scored as 0), the length of responses to indirect questions should not be strongly related to bridging and elaboration scores. For example, in the Appendix, the responses to sentence 25 are of different length and detail, but both received full points for bridging and elaboration.

Another scoring system was developed to assess the quality of responses to the direct questions (Magliano et al. 2011). R-SAT uses an "ideal" answer score each answer to each direct question. For this study, trained raters identified the amount of coverage of the ideal answers. Responses were scored on a four-point scale (0-3). A 3 indicated that the answer was complete; a 2 indicated that it was almost complete; a 1 indicated that the answer was vague, but somewhat correct; finally, a 0 indicated that the answer was incorrect. Inter-rater reliability was high (κ = .89). The Appendix contains samples of responses to direct questions. As can be seen, there is variability in the degree to which participants produced responses that reflected a coherent mental model. For example, although the participants in the appendix reflected on similar content and used bridging at the first indirect question, their responses to the first direct question showed markedly different understandings of the text.

Procedure

The current study consisted of two one-and-a-half hour sessions that were held approximately 48 hours apart. Participants completed the first session in a classroom setting equipped with individual desks. In the first session, participants were administered the Gates-McGinitie reading test, Shipley Vocabulary Test, the Author Recognition Questionnaire, the Metacognitive Strategies Index, and the Need for Cognition Scale. Participants were given a packet containing each of these measures. Participants worked through these assessments at their own pace and were dismissed individually upon completion.

In the second session, participants completed the RSAT and the OSPAN tasks. The second session was completed in a room equipped with 6 individual computer workstations with dividers between them. Participants were given the following RSAT instructions:

> Read the following texts. They are presented one sentence at a time and only one sentence will be on the screen at a time. When you have completed reading each

> sentence, click on the continue button. Periodically you will be asked to report your thoughts or answer a question and a text box will appear. Type your thoughts about the current sentence as it fits with your understanding of the text so far.

When answering the direct questions, participants were instructed to answer as completely and accurately as possible. When reporting their thoughts after seeing the prompt "What are you thinking about now?" participants were instructed to report whatever thoughts they were having about the text at the time. They were told that thoughts should be relevant to the text, but that any thoughts they may have regarding their understanding of the text should be included. Participants answered questions and reported their thoughts by typing them into a box at the bottom of the screen.

RSAT affords different text and prompt presentation options. The "one sentence at a time option" in RSAT was used in which participants only saw one sentence at a time. This was used because it requires students to draw upon their mental model when producing response to the direct and indirect prompts (Gilliam et al., 2007). However, the texts contained multiple paragraphs and in order to signal the transition to a new paragraph, "NEW PARAGRAPH" markers were presented to indicate where paragraph breaks occurred. We also used an option in which the sentence preceding a direct or indirect prompt was not available when the prompt was shown. That is, participants did not have any text available when they were responding to the think-aloud prompt.

Before starting RSAT administration, participants completed a paper and pencil practice packet. This practice required participants to read one sentence per page and to answer one direct and three indirect prompts. They wrote out their thoughts and answered the questions directly on the practice packet. The experimenters provided them with feedback prior to starting the computer version of RSAT. The feedback protocol that was used corresponded to that of Magliano et al. (2011; see also Trabasso & Magliano, 1996). Feedback was only given if the responses were short and lacked semantic content (e.g. "Ok" or "I don't know") and involved reiterating instructions. If the practice protocols were reflective of processing beyond these simple statements, feedback was general and reflected that the participants were responding in accordance with the instructions. This approach was taken to encourage participants to produce thoughts in accordance with the instructions, without influencing the nature of the inferences that were reported. Upon completing the RSAT, participants were administered the automated version of the OSPAN (Unsworth et al., 2005). This was administered on a computer and the instructions were automated. This version also included a practice application.

Results

There were three sets of analyses conducted to address the research questions addressed in this study. Descriptive statistics are presented prior to those analyses. Table 1 contains the means and standard deviations for each measure.

Table 1

Means and Standard Deviations of Variables

Variable	Mean	Standard Deviation
Vocabulary (Shipley)	26.42	4.09
General Comprehension Ability (Gates)	28.14	8.66
Reading Exposure	5.50	3.16
Working Memory Capacity	38.40	17.40
Meta-Cognition	12.23	3.74
Need for Cognition	8.59	19.24
RSAT Direct Question Score	1.31	0.44
RSAT Bridging	2.34	0.64
RSAT Elaboration	1.08	0.35

Two multivariate outliers were present in the data (based on Mahlanobis Distance measures) and were deleted from the analyses. Table 2 contains the bivariate correlations between the measures. The data were found to be multivariate skewed (Mardia's Coefficient = 1.899, C.R. = 1.376), but this value is below traditional cutoffs (e.g., Hair, Black, Babin, Anderson, & Tatham, 1998), therefore, no corrections were made.

Table 2

Correlation Matrix of Variables

Construct	Variable	2.	3.	4.	5.	6.	7.	8.	9.
Language-Specific Resources	1. Vocabulary (Shipley)	.59***	.44***	.15†	.21	.11	.56***	.25**	.27**
	2. General Comprehension Ability (Gates)		.45***	.27**	.28**	.21*	.54***	.32***	.22**
	3. Reading Exposure			.04	.12	.08	.28*	.19*	.17*
Domain-General Resources	4. Working Memory Capacity				.14	.05	.19*	.11	.01
	5. Meta-Cognition					.22**	.19*	.12	.06
	6. Need for Cognition						.21*	.19*	.03
Mental Model Construction	7. RSAT Direct Question Score							.60***	.29***
Inference Processes	8. RSAT Bridging								.15†
	9. RSAT Elaboration								

†$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$

Do measures that are theoretically domain-general or language-specific share common variance with each other?

To address this question, we conducted a principal components analysis with the relevant domain-general (operation span, need for cognition, and metacognition) and language-specific measures (Gates, vocabulary, and exposure to text) using varimax rotation. The Kaiser-Meyer-Olkin test suggested "middling" sampling of the items (KMO = .70; Kaiser, 1974). There were two components with eigenvalues greater than 1. In addition, a parallel analysis was conducted.

This analysis plots the eigenvalues of the derived components against randomly generated eigenvalues. The logic of this test is that real components should have higher eigenvalues than randomly generated data (Hayton, Allen, & Scarpello, 2004). The parallel analysis supported a two-component solution, so this solution was retained. This solution accounted for 55.73% of the variance in the items. The component structure can be seen in Table 3. As can be seen, the principal components analysis revealed that there is, in fact, shared variance among measures that represent domain-general resources and those representing language-specific ones. In short, the solution agreed with our hypothesized components, which we labeled domain-general and language-specific resources. Importantly, despite the fact that the measure of metacognitive strategies was grounded in reading, coefficients reported in Table 3 clearly show that it did not load strongly onto the language-specific component.

Table 3

Component Loadings from PCA with Varimax Rotation

Item	Language-Specific Resources	Domain-General Resources
Reading Exposure	**.80**	-.05
Vocabulary (Shipley)	**.82**	.15
General Comprehension Ability (Gates)	**.77**	.37
Need for Cognition	-.04	**.79**
Metacognition	.18	**.60**
Working Memory Capacity	.12	**.50**

Does inference generation account for any variance in comprehension above and beyond that explained by language-specific and domain-general resources?

First, assumptions of normality and multicollinearity were assessed. These assumptions were found to be met. Next, to test the research question comprehension scores were regressed onto the language-specific and domain-general component scores (derived using regression scores) in step 1, followed by elaboration and bridging scores in step 2. Step 1 accounted for a significant portion of variance in direct question response scores, $R^2 = .34$, $F(2, 141) = 36.19$, $p < .001$. Step 2 predicted variance over and above this, resulting in a significant final model, which predicted about 54% of the variance in direct question response scores, $R^2 = .54$, $F(4, 139) = 40.06$, $p < .001$. The *b*-weights and 95% confidence intervals of this analysis are presented in Table 4.

As can be seen, inference generation does appear to predict variance in the comprehension scores above basic language-specific and domain-general resources. However, this test does not offer insight into the relationship between the inference, resource variables and comprehension. Thus, the next question in the present study addressed this issue by using path analysis to explore potential mediating effects.

Table 4

Regression Model Results for Predictors After Step 2

Step	Predictor	*b*	*β*	*SE*	95% Confidence Intervals for *b*
Step 1	Domain-General Skills	.08**	.18	.03	[.03, .13]
	Language-Specific Skills	.15***	.26	.03	[.10, .21]
Step 2	Bridging Inferences	.29***	.44	.04	[.21, .37]
	Elaborative Inferences	.16*	.13	.07	[.01, .30]

*$p < .05$, ***$p < .001$

Does the data support the partially mediated model assumed by the IMM?

A path analysis was constructed which used domain-general and language-specific components as the predictor variables, elaboration and bridging scores as the mediators, and direct question response scores as the criterion variable. Data were analyzed using SPSS AMOS 23.0 (Arbuckle, 2013) using maximum likelihood estimation. This analysis compared the IMM with fully mediated (i.e., no direct relationship between resources and mental model construction) and non-mediated (i.e., no mediating paths through bridging and elaboration) nested models. Table 5 has fit indices for the three models. This analysis was conducted with 5000 bootstrapped samples.

The fit indices in Table 5 indicate that the partially mediated model was the best fitting model compared to fully and non-mediated models. However, the path from domain-general resources to elaborative inferences was nonsignificant. Therefore, this path was eliminated and the model was re-analyzed. This final model was found to have excellent fit, as noted in Table 5. Figure 2 contains the final model's standardized β-weights. The indirect effects of domain-general, $\beta = .09$, 95% CI [.01, .17], $p = .024$, and language-specific resources, $\beta = .16$, 95% CI [.08, .24], $p = .001$, significantly predicted variance in direct question response scores. In total, the final model accounted for about 53% of the variance in direct question response scores.

Table 5

Model Fit Indices for Question 3

Models	χ2	Df	Δχ2	RMSEA	90% CI for RMSEA	CFI	NFI
Partial Mediation	0.92	2		.00	[0.00, 0.13]	1.00	.99
No Mediation	29.20***	6	28.28***	.16	[0.11, 0.23]	.82	.79
Full Mediation	37.81***	4	36.89***	.24	[0.18, 0.32]	.74	.73
Partial Mediation (No Path from Domain-General to Elaborative Inferences	1.06	3	0.14	.00	[0.00, 0.09]	1.00	.99

***$p < .001$

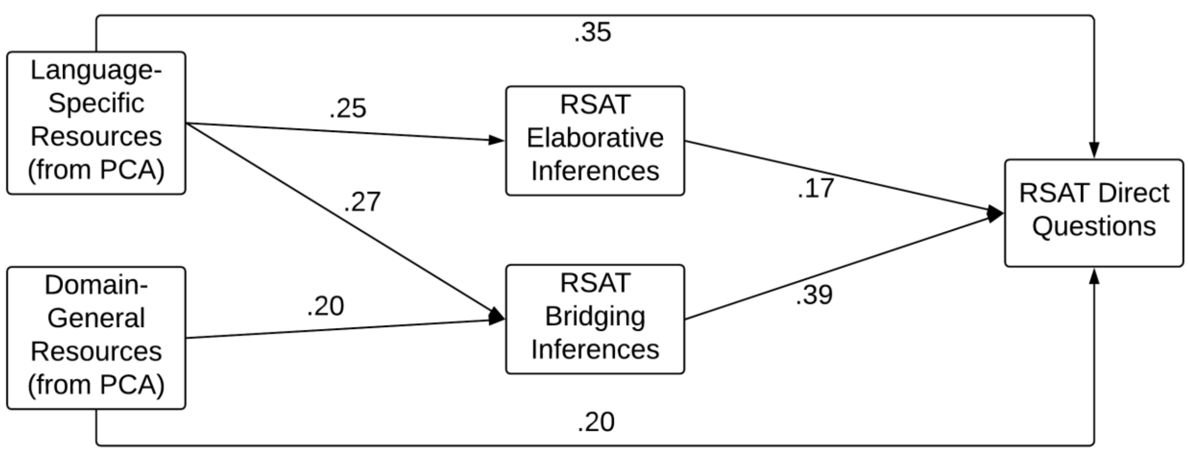

Figure 2. Final empirical model with standardized estimates. This model is based on the IMM in that we conceptualize responses to the RSAT direct questions as a measure of mental model construction. Note that the path from domain-general resources to elaborative inferences is constrained to 0.

Discussion

Readers bring a variety of resources to any given reading situation. However, there are relatively few models that explore the role that these resources play in the comprehension of text. One notable exception is the DIME model (Cromley & Azevedo, 2007), which the current study has sought to extend. Specifically, we tested whether a variant of DIME model, the IMM, could describe the relationship between reader resources, inference generation, and comprehension *while* one is reading a text.

In order to test whether the IMM applies to mental model construction and inference generation, we addressed three interrelated research questions. First, we tested whether the resources that readers bring to a reading situation can be categorized as language-specific (e.g., vocabulary) or as domain-general (e.g., working memory). The results of a principal components analysis indicated that the measures of reader resources were indeed separated into the two hypothesized categories. It is notable that this was the case, even though our measure of metacognitive strategies was specific to text. While metacognitive strategies may be specific to the type of information being processed, the propensity to monitor progress and appropriately deploy strategies may be construed as a skill that extends beyond any specific domain of information processing and learning.

Next, we asked whether inference generation predicts variance in mental model construction over and above the two categories of reader resources. The results of a regression analysis indicated that two classes of inferences (bridging and elaborative inferences) did predict variance in direct question response scores over and above language-specific and domain-general resources. This finding replicates and extends prior research showing that inferences and comprehension strategies revealed when thinking aloud are predictive of the comprehension for the texts in which the protocols where produced (e.g., Magliano & Millis, 2003; Magliano et al., 2011; Millis et al., 2006). Much of the prior research on DIME focused only on comprehension skill as measured by standardized texts, and the present study shows that this finding holds up with a more contextualized measure of comprehension. This supports the claim that reader resources alone are not sufficient to support comprehension. Rather, comprehension requires one to engage in the inference processes that support mental model construction (Graesser et al., 1994; Kintsch, 1998; 1988; van den Broek Young, Tzeng, & Linderholm, 1999).

Finally, we tested the extent to which inference processes partially mediate the relationship between reader resources and mental model construction (See Figure 1), as delineated by the IMM. The results of the path analysis demonstrated that the partial mediation model provided an excellent fit to the data, and one superior to non-mediated or full mediation models. As such, the present study demonstrates robust evidence for DIME extending to inferences and mental model construction via IMM.

Interestingly, the relationship between domain-general resources and elaborative inferences was found to be nonsignificant. This result speaks to a discrepancy in the literature about working memory capacity's relationship with bridging and elaborative inferences. As mentioned in the introduction, Fincher-Kiefer and colleagues (Fincher-Kiefer, 2001; Fincher-Kiefer & D'Agostino, 2004) have argued working memory is not as necessary for bridging inferences as it is for elaborative inferences and therefore bridging inferences are less correlated with working memory than elaborative inferences. However, we found that domain-general resources are important for constructing bridging inferences but not for constructing elaborative inferences. The finding that high-span readers tend to generate more bridging inferences and low-span readers tend to generate more elaborative inferences (Whitney et al., 1991) is consistent with our findings. However, because we were not looking at group differences, the current study does not allow us to explore the nature of the relationship between working memory and bridging and elaborative inferences. There are several reasons why the discrepancy occurred, including different tasks, texts and operational definitions of elaborations and bridges. However, the relationship between domain-general resources and bridging inferences and the lack of relationship between domain-general resources and elaborative inferences supports the idea that bridging inferences are more computationally demanding whereas elaborative inferences reflect the passive activation of long-term memory stores (Kintsch, 1998; McKoon & Ratcliff, 1992; Myers & O'Brien, 1998). More research is needed to clarify the relation between working memory resources and types of inferences that are generated during comprehension.

As discussed in the introduction, the DIME model attempts to explain general comprehension proficiency (Ahmed et al., 2016; Cromley & Azevedo, 2007). The present study demonstrates that the central idea of DIME—reader resources directly support comprehension and indirectly support it through inference ability—also applies to comprehension as measured by the quality of one's mental model (Kintsch, 1988; Zwaan & Radvansky, 1998). The distinction between comprehension and mental model construction is similar to one made by Zwaan and Radvansky (1998) in the context of the event indexing model, a prominent theory of mental model construction. That is, they make a distinction between an *integrated model* and *complete model* (see also Langston, Trabasso, & Magliano, 1999). An integrated model reflects the status of the mental model after each sentence is read and the mental model is updated to incorporate that content. The complete model reflects the nature of the mental model after all sentences have been read and the mental representation becomes stabilized (i.e., no longer updated based on reading new content). Importantly, the current study brings DIME more in line with most other models of comprehension (e.g., Gernsbacher, 1990; Graesser et al., 1994; Kintsch, 1988) in that it focuses on comprehension as a process rather than a static skill.

While it is normally best to measure independent and dependent variables with separate measures, our research questions center on how resources influence the inference processes that occur while reading a particular text and the effect of these inference processes on the comprehension of those same texts. Thus, the current study required an instrument, such as RSAT, that afforded the ability to measure inferences and comprehension within the same text. There is precedence in think-aloud studies to assess the relationship between processes revealed in the think aloud protocols and the comprehension of those texts (e.g., Magliano & Millis, 2003). While it is important to acknowledge this aspect of the study, it is also important to emphasize that the research questions did not focus on the relative importance of reader resources and inference processes on comprehension. Moreover, the partially mediated IMM replicated the findings of prior tests of the DIME model in that inferences mediated the relationship between reader resources and comprehension (Ahmed et al., 2016, Cromley & Azevedo, 2007; Cromley et al., 2010). As such, this aspect of the study does not compromise our ability to answer the central research questions.

As is common in research using thinking aloud to study inference processing, the data in the present study were aggregated across think aloud location (e.g., Magliano & Millis, 2003). Doing so was necessary to conduct the analyses reported here. However, one does not have to rely on aggregated thinking aloud data (e.g., Coté & Goldman, 1999; Pressley & Afflerbach, 1995). This latter approach affords tracing participants' reasoning and thoughts as they incrementally build a mental model for the text. Moreover, one could attempt to account for how the text guides the reader's thoughts across time by inspecting how reader's thoughts align with the content of the text (Allen, Jacovina, & McNamara, 2016; Allen, Snow, & McNamara, 2015). It might even be possible to examine how different resources could manifest themselves in the protocols (Allen et al., 2015). For example, limitations of working memory might be observed if a reader fails to mention earlier text when a model of the text specifies that it should be accessible. To do this, however, one would need at a minimum to collect protocols after each sentence (or even clause) to get a complete landscape of the reader's evolving mental model. Given the number of texts that were read, and the fact that participants also answered comprehension question while reading the text, doing so was not feasible in the present study. Nonetheless, exploring the dynamic nature of mental model constructs and how reader resources affect them over the course of reading a text certainly has merits for future research.

Although a particular reader may have a set of resources available to deploy during reading, he or she will not uniformly use such resources for all texts. Readers likely will have different levels of engagement for different texts based on what an acceptable level of coherence is for a given reading situation (e.g., Graesser et al., 1994; van den Broek et al., 2011; van den Broek et al., 2001). Although the IMM does not yet account for differing engagement levels, both it and the DIME are consistent with the idea of standards of comprehension playing a role in comprehension, because they underscore the idea that reading is a complex system of resources and processes which are highly dependent on context. It is a valuable endeavor to isolate the components of comprehension as has been previously done (e.g., Beck et al., 1982; Cunningham & Stanovich, 1991; Daneman & Carpenter, 1980). However, we believe it is also important to consider comprehension as a larger system of resources and processes that work together to form a mental model of a particular text.

Studying comprehension through the lens of a system of interconnected resources and processes has key benefits. For example, if researchers can understand how different readers approach different texts, more individualized interventions can be created to assist students in reading. For example, McMaster et al., (2012) found that interventions did not differ in their efficacy across reading skill level, but they did so when poor readers were separated into subgroups based on their think-aloud protocols. Based on findings such as this, we believe that the current study offers a valuable addition to the literature on how reader and situational differences can affect comprehension and may influence more targeted interventions for reading difficulties.

The current study was conducted with a college-aged sample, and previous work on DIME has demonstrated that it applies to primary and secondary school samples (Ahmed et al., 2016; Cromley & Azevedo, 2007), so we believe that the IMM should also apply to younger samples. However, confirmatory work is needed. Additionally, although the current study expanded the scope of DIME in regards to certain resources (e.g., need for cognition), there are other resources relevant to comprehension that need to be explored in the context of contextualized inference generation and mental model construction such as prior knowledge (Cromley & Azevedo, 2007; Kendeou & van den Broek, 2007; McNamara & Kintsch, 1996). Finally, future research should address how mental model construction occurs across different genres of text within the context of IMM. As has been shown in prior work, there are key differences between different genres of text and the process of comprehension changes accordingly (e.g., narrative, science, and. history; Otero, León, & Graesser, 2000), so an important question is how comprehension occurs in each of these genres. Unfortunately, answering this question falls outside of the scope of the current study's data, so future work will need to address such questions in the IMM framework using think-aloud data.

It should be noted that no measure of writing skill or typing proficiency was given to participants. Despite previous research that shows that typing-aloud and thinking-aloud produce similar results (Muñoz, Magliano, Sheridan, & McNamara, 2006), it is possible that there are differences in the results produced in the current study's methodology compared to a more traditional think-aloud procedure.

In conclusion, the current study provides an understanding of how comprehension emerges from a complex system of resources and processes rather than existing merely as a static skill that readers possess. While research that focuses on specific parts of this system has been valuable, the current study illustrates the importance of designing studies that can explore the relationships between what readers bring to a reading situation and what they actually do in that situation.

References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50-68. doi: 10.1037/a0014411

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.

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12.*Contemporary Educational Psychology*, *44*, 68-82. doi: 10.1016/j.cedpsych.2016.02.002

Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). Cohesive features of deep text comprehension processes. In J. Trueswell, A. Papafragou, D. Grodner, & D. Mirman (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society in Philadelphia, PA*, (pp. 2681-2686). Austin, TX: Cognitive Science Society.

Allen, L. K., Snow, E. L., & McNamara, D. S. (2015). Are you reading my mind? Modeling students' reading comprehension skills with Natural Language Processing techniques. In J. Baron, G. Lynch, N. Maziarz, P. Blikstein, A. Merceron, & G. Siemens (Eds.), *Proceedings of the 5th International Learning Analytics & Knowledge Conference (LAK'15*), (pp. 246- 254). Poughkeepsie, NY: ACM.

Arbuckle, J. L. (2013). AMOS (Version 7.0) Chicago. *IL: SPSS.*

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506-521.

Budd, D., Whitney, P., & Turley, K. J. (1995). Individual differences in working memory strategies for reading expository text. *Memory & Cognition*, *23*(6), 735-748. doi: 10.3758/BF03200926

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197-253.

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. doi: 10.1037/0022-0663.96.1.31

Calvo, M. G. (2005). Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading. *Learning and Individual Differences*, *15*(1), 53–65. doi: 10.1016/j.lindif.2004.07.002

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory*, *9*(4-6), 365-381. doi: 10.1080/09658210143000083

Colzato, L. S., Spapé, M. M., Pannebakker, M. M., & Hommel, B. (2007). Working memory and the attentional blink: Blink size is predicted by individual differences in operation span. *Psychonomic Bulletin & Review*, *14*(6), 1051-1057. doi: 10.3758/BF03193090

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. *The construction of mental representations during reading*, 169-193.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183. doi: 10.1016/S0160-2896(01)00096-4

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. doi: 10.1037/0022-0663.99.2.311

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*,*102*(3), 687-700. doi: 10.1037/a0019452

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.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264–274. doi: 10.1037/0022-0663.83.2.264

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. doi: 10.1207/s1532799xssr1003_5

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. doi: 10.1016/j.cedpsych.2006.05.002

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466. doi: 10.1016/S0022-5371(80)90312-6

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, *130*(2), 208-223. doi: 10.1037/0096-3445.130.2.208

Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: The role of distraction control. *Neuropsychology*, *22*(5), 638-644. doi: 10.1037/0894-4105.22.5.638

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287. doi: 10.1598/RRQ.36.3.2

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. doi: 10.1111/1467-8721.00160

Farley, M. J., & Elmore, P. B. (1992). The relationship of reading comprehension to critical thinking skills, cognitive ability, and vocabulary for a sample of underachieving college freshmen. *Educational and Psychological Measurement*, *52*(4), 921–931. doi: 10.1177/0013164492052004014

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*(4), 541–553. doi: 10.1016/j.jml.2005.12.006

Fincher-Kiefer, R. (1996). Encoding differences between bridging and predictive inferences. *Discourse Processes, 22*(3), 225-246, DOI: 10.1080/01638539609544974

Fincher-Kiefer, R. & D'Agostino, P. R. (2004). The role of visuospatial resources in generating predictive and bridging inferences. *Discourse Processes*, *37*(3), 205-224. doi: 10.1207/s15326950dp3703_2

Fox, E. & Riconscente, M. (2008). Metacognition and self-regulation in James, Piaget, and Vygotsky. *Educational Psychology Review 20*(4), 373-389. doi: 10.1007/s10648-008-9079-2

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Psychology Press.

Gilliam, S., Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2007). Assessing the format of the presentation of text in developing a Reading Strategy Assessment Tool (R-SAT). *Behavior Research Methods*, *39*(2), 199–204. doi: 10.3758/BF03193148

Graesser, A. C., & Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension.*Advances in Psychology*, *29*, 53-94. doi: 10.1016/S0166-4115(08)62732-6

Graesser, A. C., & Franklin, S. P. (1990). QUEST: A cognitive model of question answering. *Discourse Processes*, *13*(3), 279-303. doi: 10.1080/01638539009544760

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. doi: 10.1037/0033-295X.101.3.371

Guthrie, J. T., McGough, K., Bennett, L., & Rice, M. E. (1996). Concept-oriented reading instruction: An integrated curriculum to develop motivations and strategies for reading. *Developing Engaged Readers in School and Home Communities*, 165–190.

Hacker, D. J. (1998). Definitions and empirical foundations.*Metacognition in educational theory and practice*, *14*, 1-23.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis* (Vol. 5, No. 3, pp. 207-219). Upper Saddle River, NJ: Prentice hall.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods*, *7*(2), 191-205. doi: 10.1177/1094428104263675

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. doi: 10.1037/0033-295X.99.1.122

Kaiser, H. F. (1974). An index of factor simplicity. *Psychometrika, 39*(1), 31-36.

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, *35*(7), 1567–1577. doi: 10.3758/BF03193491

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310-321. doi: 10.1037/0022-0663.100.2.310

Kurby, C. A., Britt, M. A., & Magliano, J. P. (2005). The role of top-down and bottom-up processes in between-text integration. *Reading Psychology, 26,* 335-362. doi: 10.1080/02702710500285870

Langston, M. C., Trabasso, T., & Magliano, J. P. (1999). A connectionist model of narrative comprehension. In *Understanding language understanding* (pp. 181-226). MIT Press.

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints.*Discourse Processes*, *34*(3), 259-280. doi: 10.1207/S15326950DP3403_2

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, *94*(4), 778-784. doi:10.1037/0022-0663.94.4.778

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities.*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1424-1429. doi: 10.1037/0278-7393.27.6.1424

Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers' sentence-and text-level representations. *Journal of Memory and Language*, *36*(1), 129-145. doi: 10.1006/jmla.1996.2485

Long, D. L., Seely, M. R., Oppy, B. J., & Golding, J. M. (1996). The role of inferential processing in reading ability. In *Models of understanding text*, 189-214.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*(1), 35-54. doi: 10.1037/0033-295X.109.1.35

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98. http://doi.org/10.1016/0010-0285(92)90003-K

MacGinitie, W. H. (2000). *Gates-MacGinitie reading tests*. (Vol. levels 7/9 & 10/12; forms S&T). *Itasca, IL: Riverside Publishing*.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *20*(3), 193-232. doi: 10.1016/0304-422X(91)90007-C

Magliano, J. P., Larson, A. M., Higgs, K., & Loschky, L. C. (2016). The relative roles of visuospatial and linguistic working memory systems in generating inferences during visual narrative comprehension. *Memory & cognition, 44*(2), 207-219. doi: 10.3758/s13421-015-0558-7

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *21*(3), 251–283. doi: 10.1207/S1532690XCI2103_02

Magliano, J. P., Millis, K. K., The RSAT Development Team, & Levinstein, I.,. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*(2), 131–154. doi: 10.1007/s11409-010-9064-2

McBride-Chang, C., Manis, F. R., Seidenberg, M. S., Custodio, R. G., & Doi, L. M. (1993). Print exposure as a predictor of word reading and reading comprehension in disabled and nondisabled readers. *Journal of Educational Psychology*, *85*(2), 230-238. doi: 10.1037/0022-0663.85.2.230

McKoon, G. & Ratcliff, R. (1992). Inferences during reading. *Psychological Review, 99*(3), 440-466.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 82-91. doi: 10.1037/0278-7393.12.1.82

McMaster, K. L., Van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., ... & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*,*22*(1), 100-111. doi: 10.1016/j.lindif.2011.11.017

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse processes*, *38*(1), 1-30. doi: 10.1207/s15326950dp3801_1

McNamara, D. S., de Vega, M., & O'Reilly, T. (2007). Comprehension skill, inference making, and the role of knowledge. In *Higher level language processes in the brain: Inference and comprehension processes*, 233-251.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, *22*(3), 247-288. doi: 10.1080/01638539609544975

McNamara, D. S., & Magliano, J. P. (2009a). Self-explanation and metacognition: The dynamics of reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser, (Eds.), *Handbook of Metacognition in Education* (pp. 60-81). Mahwah, NJ: Lawrence Erlbaum and Associates.

McNamara, D. S., & Magliano, J. P. (2009b). Towards a comprehensive model of comprehension. In B. Ross (Ed), *The Psychology of Learning and Motivation*, vol. 51, (pp. 297-384). New York, NY, US: Elsevier Science.

McNamara, D. S., O'Reilly, T. P., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*(2), 147-171. doi: 10.2190/1RU5-HDTJ-A5C8-JVWE

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*(2), 302–320. doi: 10.1037/a0025250

Medo, M. A., & Ryder, R. J. (1993). The effects of vocabulary instruction on readers' ability to make causal connections. *Literacy Research and Instruction*, *33*(2), 119-134. doi: 10.1080/19388079309558149

Millis, K. & Magliano, J. (2012). Assessing comprehension processes during reading. In J. Sabatini, E. Albro, & T. O'Reilly (Eds.) *Reaching an Understanding: Innovations in How We View Reading Assessment. (pp. 35-53).* R&L Education.

Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. *Scientific Studies of Reading*, *10*(3), 225–240. doi: 10.1207/s1532799xssr1003_2

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, *94*(2), 249-259. doi: 10.1037/0022-0663.94.2.249

Muñoz, B., Magliano, J. P., Sheridan, R., & McNamara, D. S. (2006). Typing versus thinking aloud when reading: Implications for computer-based assessment and training tools. *Behavior Research Methods, 38*(2), 211-216. doi: 10.3758/BF03192771

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

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. *Vocabulary Acquisition: Implications for Reading Comprehension*, 52–77.

National Center for Education Statistics. (2015). *The nation's report card: 2015 mathematics and reading assessments* (NCES Report No. 2015136). Washington, D. C.: National Center for Education Statistics.

Nixon, S. J., Parsons, O. A., Schaeffer, K. W., & Hale, R. L. (1995). A methodological study of the Shipley Institute of Living Scale in alcoholics and non-alcoholics: reliability, discriminating items and alternate forms. *Applied Neuropsychology*, *2*(3/4), 155.

Oberauer, K., Süß, H. M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*(2), 167-193. doi: 10.1016/S0160-2896(02)00115-0

O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, *43*(2), 121-152. doi: 10.1080/01638530709336895

Otero, J. C., León, J. A., & Graesser, A. C. (2000). *The psychology of the scientific* *text*. Erlbaum: New York.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of educational psychology*, *98*(3), 554-566. doi: 10.1037/0022-0663.98.3.554

Paris, S. G., Wasik, B., & Turner, J. C. (1991). The development of strategic readers. Erlbaum: New York.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational psychology*, *67*(4), 461-469. doi: 10.1037/h0077013

Perfetti, C., McKeown, M. G., & Kucan, L. (2010). Decoding, vocabulary, and comprehension. *Bringing reading research to life*, 291-303.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22-37. doi: 10.1080/10888438.2013.827687

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*, 289–312. doi: 10.1080/10888430701530417

Schmitt, M. C. (1988). The effects of an elaborated directed activity on the metacomprehension skills of third graders. *Dialogues in Literacy Research*, 167–181.

Schmitt, M. C. (1990). A questionnaire to measure children’s awareness of strategic reading processes. *The Reading Teacher*, *43*(7), 454–461.

Singer, M. (1988). Inferences in reading comprehension. In M. Daneman, G. E. Mackinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice,* Vol. 6, pp. 177-219). San Diego, CA: Academic Press.

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, *20*(5), 539–548. doi: 10.3758/BF03199586

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, *85*(2), 211.

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

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In *Handbook of metacognition in education*, 85-106.

Todaro, S.A., Magliano, J.P., Millis, K.K., McNamara, D.S., & Kurby, C.A. (2008). Assessing the structure of verbal protocols. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 607-612). Austin, TX: Cognitive Science Society.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(3), 255–287. doi: 10.1080/01638539609544959

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*(1-2), 3-34. doi: 10.1080/01638539309544827

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. doi: 10.1016/0749-596X(89)90040-5

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. doi: 10.3758/BF03192720

van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-588). San Diego, CA: Academic Press.

van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123-139). Charlotte, NC: IAP 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. doi: 10.3758/BF03206376

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. *The Construction of Mental Representations during Reading*, 71-98.

Vasquez, B. P., Binns, M. A., & Anderson, N. D. (2017). Response time consistency is an indicator of executive control rather than global cognitive ability. *Journal of the International Neuropsychological Society*, 1-10. doi: 10.1017/S1355617717001266

Virtue, S., Parrish, T., & Jung-Beeman, M. (2008). Inferences during story comprehension: Cortical recruitment affected by predictability of events and working memory capacity. *Journal of Cognitive Neuroscience*, *20*(12), 2274-2284.doi: 10.1162/jocn.2008.20160

Was, C. A., Rawson, K. A., Bailey, H., & Dunlosky, J. (2011). Content-embedded tasks beat complex span for predicting comprehension. *Behavior Research Methods*, *43*(4), 910-915. doi: 10.3758/s13428-011-0112-x

Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse processes*, *14*(2), 133-145. doi: 10.1080/01638539109544779

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2007). *Development of achievement motivation*. John Wiley & Sons, Inc.

Zachary, R. A., & Shipley, W. C. (1986). *Shipley institute of living scale: Revised manual*. WPS, Western Psychological Services.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*(5), 292-297. doi: 10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162-185. doi: 10.1037/0033-2909.123.2.162