THE EFFECTS OF READER CHARACTERISTICS, TEXT FEATURES, AND COMPREHENSION PROCESSES ON READING COMPREHENSION

A Dissertation

Presented to

The Faculty of the Department

of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Paulina A. Kulesz

May, 2014

THE EFFECTS OF READER CHARACTERISTICS, TEXT FEATURES, COMPREHENSION PROCESSES ON READING COMPREHENSION

Paulina A. Kulesz

APPROVED:

David J. Francis, Ph.D.
Committee Chair

Jack M. Fletcher, Ph.D.

Jennifer Tackett, Ph.D.

Marcia Barnes, Ph.D.
University of Texas at Austin

John W. Roberts, Ph.D.
Dean, College of Liberal Arts and Social Sciences
Department of English

THE EFFECTS OF READER CHARACTERISTICS, TEXT FEATURES, COMPREHENSION PROCESSES ON READING COMPREHENSION

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Paulina A. Kulesz

May, 2014

ABSTRACT

When the psychological processes underlying test performance are understood, psychometric theory dictates that item difficulty can be explained through these processes. The goal of the project was to improve understanding of reading comprehension and the role of reader characteristics, passage features, and comprehension processes in understanding text through the application of explanatory item response models. Participants were 1,190 students from 11 to 20 years recruited from eight schools within four districts. Students represented a range of reading comprehension skills in terms of their word reading, semantic language and operational (i.e., inference making, working memory capacity) skills. Measures included the grade appropriate Gates-MacGinitie reading comprehension and vocabulary subtests, letter word identification and numbers reversed subtests of the Woodcock & Johnson tests of cognitive abilities, test of word reading efficiency, as well as a researcher-developed test of background knowledge. The results indicated that reader characteristics including vocabulary, background knowledge, working memory and reading fluency were the most influential in explaining variation in reading comprehension item performance. Passage features explained some variation in item difficulties, with expository passages and deep cohesion being the most influential. Most importantly, a few text-reader interactions affected reading comprehension test scores. However, their effects were not pronounced, as good readers tended to perform better than poor readers regardless of the text they read. Better word and world knowledge was found to be the most helpful in understanding texts of variable difficulty. These findings are consistent with research that targets the building of vocabulary skill and background knowledge in order to improve reading comprehension, and suggests that the benefits of such development would apply to a wide variety of texts and to both memory for what has been read as well as drawing inferences from the text. The study further showed that explanatory item response models can be applied in a meaningful way to operational standardized tests, while also highlighting the limitations inherent in such application for explicating the general effects of text characteristics and reader abilities on the comprehension of written language.

Keywords: reading comprehension, reader characteristics, text features, comprehension processes, explanatory item response models

ACKNOWLEDGEMENTS

I would like to express special thanks to my advisor, Dr. David J. Francis for his great guidance, commitment and interest in the pursuit of the study, and preparation of this dissertation.

To Drs. Jack M. Fletcher, Marcia Barnes and Jennifer Tackett, I would like to express my gratitude for their guidance in expanding my knowledge about topics concerning this dissertation.

Table of Contents

List of Tables

DEDICATION

I dedicate this dissertation to my beloved parents and husband who always support me on the way to success.

The Effects of Reader Characteristics, Text Features, and Comprehension Processes on Reading Comprehension

Reading comprehension is about understanding of language presented in the written form or text. Historically, definitions of reading comprehension have focused on lower-level processes including word recognition, decoding, phonological awareness, and grammatical structure. As reading research advanced it became clear that reading comprehension cannot be simply defined in terms of lower-level processes and their interactions, as some readers have problems with comprehending the text despite adequate lower-level reading skills. This led to the conclusion that lower-level processes are necessary but not sufficient for becoming a proficient reader, which is captured elegantly by the Simple View of Reading (SVR; Gough & Tunmer, 1986). SVR posits that reading comprehension is the product of decoding skills and listening comprehension skills, such that smaller values for either skill will result in reduced reading comprehension. More recently, reading comprehension research has focused on deeper levels of comprehension. The focus of the research shifted towards ability to create a coherent, mental representation of the text. Many researchers started acknowledging the fact that both lower-level and deeper level processes are important aspects of reading comprehension.

At present, many models, such as the Construction-Integration and Landscape models discussed in more detail below, describe reading comprehension as a complex, cognitive phenomenon involving remembering the text, understanding main concepts in the text, and constructing mental representations of the text (Snow & Polselli Sweet, 2003). This phenomenon is embedded in a sociocultural context and depends on dynamic relations between reader, text, and cognitive processes occurring as the reader progresses through the text (Snow & Polselli Sweet, 2003). Despite significant progress in studying effects of text, reader, and cognitive processes on reading comprehension, interactions between these elements and their effects on comprehension remain understudied. The present study proposes to examine these interactions as their careful investigation carries practical implications for intervention and assessment.

The introduction of the manuscript is divided into four sections. To begin, the comprehension processes including description of the construction-integration and Landscape models as well as distinction between process and product of reading comprehension are discussed. This section is followed by an overview of reader characteristics and text features affecting reading comprehension. Finally, explanatory item response models for binary data are reviewed.

Comprehension Processes: Cognitive Processes involved in Reading

Reading comprehension, as a complex activity/cognitive task involving information processing, is vastly dependent on cognitive processes which take place as the reader progresses through the text (Hannon & Daneman, 2001; Kintsch, 1979; Rapp & van den Broek, 2005; Yeari & van den Broek, 2011). These cognitive processes, such as remembering text, making inferences from the text, and integrating information from the text with existing knowledge are known collectively as comprehension processes. Comprehension processes are typically associated with deeper levels of comprehension and are crucial for successful understanding of the text (Kintsch, 1988; van den Broek, van den Broek, Rapp, & Kendeou, 2005). Depending on the information available to the reader within each reading cycle, which is typically a sentence, these processes are employed to update and integrate new information with existing textual information and/or background knowledge in order to construct a mental representation of the text (Kintsch, 2005; Lynch & van den Broek, 2007; Tzeng, van den Broek, Kendeou, & Chengyuan, 2005; van den Broek et al., 2005). A few models have been proposed to address the importance of deeper level processes in reading comprehension. These models primarily concentrate on the aspects of comprehension which take place once the words within the text are decoded/recognized. Among the most notable models are the construction-integration (Kintsch, 1988) and Landscape models (van den Broek, Young, Tzeng, & Linderholm, 1999). Both of these models are based on the notion of network representations involving meaningful relations between informational units (nodes) and semantic relations between informational units and/or background knowledge (connections; Kintsch, 1988; van den Broek et al., 1999).

The construction-integration model. The construction-integration model explains how reader interaction with the text leads to the development of a mental representation of that text. This model assumes that within each reading cycle understanding of information currently processed depends on information presented earlier in the text and/or relevant background knowledge (Kintsch, 1988). These two sources of information are crucial for creating a coherent, mental representation of the text involving three interacting levels: (a) surface structure, (b) textbase model, and (c) situation model. The linguistic surface structure of the text associated with lower-level comprehension processes represents actual words and syntactic relations occurring between these words (Kintsch, 1988). Deeper comprehension processes take place at the levels of the textbase and situation models. These levels are subsequent to linguistic recognition of the text, and describe more advanced text processing associated with construction and integration phases of the construction-integration model (respectively).

The textbase model represents information presented in the text, and might be thought of in terms of informational units (propositions/concepts) used to form the textbase representation of the text. A good textbase representation provides a clear meaning of the text and its structure. The textbase representation is an initial representation which only involves textual information. This representation is sloppy and incoherent as it is not related to a broader context associated with the reader's world knowledge. Construction of the textbase representation is based on the reader's vocabulary knowledge (knowledge necessary to identify words and relate them to other words). This knowledge is required to from textbase representation of concepts presented in the surface code (Kintsch, 1988). In order to construct the textbase representation, the reader needs to recall information from the text and make text-based bridging inferences between informational units presented in the text.

The situation model characterizes the reader's representation of the text on the basis of textual information and world knowledge. This model might be understood as a process of integrating background knowledge with the text to update a constructed representation of the text. Consequently, the situation model represents the process of reshaping the textbase representation and associating this representation with world knowledge to create a coherent representation of the text (Ericsson & Kintsch, 1995). Construction of the situation model depends on the reader's individual goals and one's level of expertise associated with relevant world knowledge. Readers who possess superficial knowledge concerning a topic introduced in the text oftentimes fail to create the situation model despite having the ability to construct the textbase representation (Kintsch, 1988). For these readers integrating text information with background knowledge is effortful as they cannot easily retrieve the relevant knowledge from their long-term memory. Consequently, they experience problems with constructing a coherent mental representation of the text, which negatively affects their comprehension. In order to compensate for gaps in their knowledge they are forced to continuously search their memory for relevant information and monitor comprehension progress in constructing the situation model (Kintsch, 2012). Not surprisingly, relevant world knowledge along with deeper comprehension processes involving knowledge integration and comprehension monitoring are important assets for readers to construct a coherent representation of the text.

The Landscape model. The Landscape model, similar to the construction-integration model, focuses on the process of constructing a mental representation of the text on the basis of the textual information and background knowledge of the reader utilizing text memory, inference and knowledge integration processes. This model not only provides a description of comprehension processes which take place during reading of the text, but also explains the role of these processes in creating a coherent, mental representation of the text (van den Broek, Rapp, & Kendeou, 2005).

Within each reading cycle, concepts presented in the text change activation status as some concepts become newly activated, others remain active or are reactivated, and yet others decline in their activation (Rapp & van den Broek et al., 2005). Newly activated concepts become integrated with the representation of the text constructed thus far whereas reactivated concepts become better remembered as they cyclically reappear during reading. If two or more concepts are co-activated within a reading cycle a connection between these concepts is established. This connection constitutes a foundation for creating associative networks between concepts. The fluctuation of activation depends on four sources of information: (a) information currently processed (reading a new sentence), (b) remaining information from the previous reading cycle, (c) representation of the text processed thus far (current episodic representation of the text based on information acquired from previous reading cycles), and (d) background knowledge (van den Broek et al., 2005; van den Broek & Kendeou, 2008). Accessing the last two sources of information depends on two mechanisms, cohort activation and coherence-based retrieval.

Cohort activation is a memory-based process that relates information currently processed with relevant textual information processed in preceding reading cycles. This process is automatic and is not controlled by the reader (van den Broek et al., 2005). Coherence-based retrieval is a process of retrieving background knowledge or episodic representation of the text constructed so far in order to meet reader's standards of coherence or goals (Linderholm, Virtue, Tzeng, & van den Broek, 2004; van den Broek, 2010). This process may be effortful or automatic. Specifically, if the reader needs to search memory to access relevant information some strategic processes might be necessary. However, in attempting to maintain local coherence the process may be automatic.

Depending on the reader's standards of coherence, ability to make inferences, memory and attentional resources, comprehension goals as well as text type, cohort activation may or may not be adequate to meet the reader's comprehension standards (Rapp & van den Broek, 2005). If the information processed within a reading cycle is not sufficiently coherent, the reader might employ coherence-based retrieval in order to improve understanding of the processed information. Consequently, activation of information through cohort activation or coherence-based retrieval leads to the construction of a mental representation of the text.

Process and product of reading comprehension. Discussion of reading comprehension requires that a distinction be made between the processes and products of reading comprehension. This distinction is clearly articulated in the construction-integration and Landscape models. In both of these models, the product of reading comprehension is defined as a mental representation of the text constructed after the completion of reading. This representation is an outcome of comprehension processes which take place as the reader progresses through the text, and is dependent on characteristics of the reader and features of the text. Successful understanding of the text results in an integrated (i.e., coherent) representation of the text involving referential and causal/logical relations of concepts presented in the text (van den Broek & Espin, 2012). These relations are frequently constructed through inferences that are formed based on textual information and/or background knowledge. The product of reading comprehension allows the reader to interact further with the text including retrieval of information, question answering, and summarization.

The comprehension processes are cognitive operations involved in text processing that determine a reader's success or failure in reading comprehension (Lorch & van den Broek, 1997). Creating a coherent mental representation of the text and employment of comprehension processes depend on limited attentional as well as working memory resources. Within each reading cycle, the reader is able to process only a limited amount of information, and to use only a subset of the available comprehension processes (Ericsson & Kintsch, 1995; van den Broek & Espin, 2012). Implementation of relevant comprehension processes determines the level of coherence of text representation. If one has not mastered lower-level comprehension skills, deeper level comprehension processes are compromised as working memory resources are used for more basic linguistic processing (Tilstra, McMaster, van den Broek, Kendeou, & Rapp, 2009).

Reader Characteristics

Reader characteristics can be understood as an umbrella term entailing abilities, skills, knowledge and experiences of the reader used to comprehend the text. These characteristics include cognitive (attention, memory, reasoning), motivational (purpose of reading, interest in the read text, and reader's self-efficacy), linguistic (word recognition, word reading/decoding, phonological awareness, understanding of grammatical structures, and inference making) skills as well as knowledge in several areas, including vocabulary, world knowledge, domain-specific knowledge, and knowledge of discourse (Snow & Polselli Sweet, 2003). Reader characteristics are important sources of variation in reading comprehension as good readers typically have a wider range of abilities and skills when compared to poor readers (Ericsson & Kintsch, 1995; Kintsch, 1994). Use of these characteristics fluctuates depending on the text type and engagement of comprehension processes. A detailed description of all relevant reader characteristics is beyond the scope of the present study. The current project focused on accurate and fluent word reading, working memory, vocabulary and background knowledge.

Word reading. Successful reading comprehension is dependent on speed and accuracy in word reading as this ability allows identifying strings of letters as specific words (Fletcher, Lyon, Fuchs, & Barnes, 2007). Fluent word reading warrants meaningful and efficient processing of words and sentences temporarily stored in working memory. Fluency in word reading not only is a prerequisite of reading comprehension but is also critical for higher level cognitive processes involved in text processing. Deficits in word reading are significant sources of variation in reading comprehension among less-skilled and beginning readers (Vellutino, 2003). Readers with inaccurate and/or labored decoding skills are unsuccessful in comprehending the text as they experience problems with reading words in the text. More precisely, their inability to interpret ideas in the text results from problems with reading the text. Allocating limited working memory resources to word reading hinders engagement of deeper comprehension processes necessary for creating a representation of the text. In other words, greater allocation of working memory resources to lower level processes limits deployment of deeper comprehension processes (van den Broek, 2012).

Working memory. Working memory capacity might be defined as operational capacity (Baddeley, 1998). This ability is important for reading comprehension as it is involved in storing and manipulating currently processed information, making inferences, and integrating text information with background knowledge in order to create a coherent representation of the text (Daneman & Hannon, 2001; Hannon, 2012). Working memory is an important factor affecting word reading. Successful word reading involves active maintenance of currently processed information and accessing stored orthographic representations of words (Christopher et al., 2012). Consequently, readers with high working memory capacity not only are better at decoding the text but also have fewer problems with integrating concepts when compared to readers with low working memory capacity (Cantor & Engle, 1993; Linderholm & van den Broek, 2002). Furthermore, readers with high capacity make more bridging inferences between information presented in the text as well as inferences between textual information and background knowledge (van den Broek, 2012).

Vocabulary and background knowledge. Vocabulary (word knowledge) and background knowledge (domain-specific and/or world knowledge) are important sources of variation among readers as they are crucial components in creating a coherent mental representation of the text. In reading comprehension, word and world knowledge closely correspond with each other as depth and breadth of reader's vocabulary can be interpreted as the linguistic equivalent of one's world knowledge (Perfetti & Adolf, 2012). Vocabulary knowledge is one of the strongest predictors of reading comprehension because word knowledge is necessary for understanding relations between words as well as deriving the meaning of sentences and concepts in the text. Consequently, readers who possess better knowledge of word forms and meanings (lexical quality) as well as know more words are more successful in comprehending the text when compared with readers' whose depth and breadth of vocabulary is more limited (Perfetti, 2007).

Background knowledge can be defined as an umbrella term entailing domain-specific and world knowledge. Domain-specific knowledge refers to knowledge pertaining to a specific content area whereas world knowledge is interpreted as the pre-existing knowledge about events and activities in everyday life relevant to the text content. Domain-specific and world knowledge are fundamental for reading comprehension as they help in understanding information conveyed in the text, integrating this information with preexisting knowledge, and filling in informational blanks in less cohesive texts (Hannon & Daneman, 1998; Kintsch, 1998; Ozuru, Dempsey, & McNamara, 2009). Relevant world knowledge is especially helpful in comprehending narrative text whereas domain-specific knowledge is particularly useful in understanding expository/informational texts (Vellutino, 2003).

Background knowledge can be activated automatically (through relation with information presented in the text) or strategically (through memory search for the relevant background knowledge). Lack of relevant background knowledge not only hinders text comprehension but also affects memory of the text (Kendeou & van den Broek, 2007). Readers who possess high background knowledge are more successful in comprehending texts because they are faster and better in accessing preexisting knowledge and integrating textual information with this knowledge when compared to low background knowledge readers (Ericsson & Kintsch, 1995; Kendeou & van den Broek, 2007; van den Broek & Kendeou, 2008; van den Broek, 2010).

Developmental age. Research suggests that working memory capacity (Conklin, Luciana, Hooper, & Yarger, 2007; Gathercole & Alloway, 2004; Hitch, Towse, & Hutton, 2001; Kemps, de Rammelaere, & Desmet, 2000), comprehension processes (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Hannon & Daneman, 2009), and the ability to understand more difficult texts (McNamara, Graesser, & Louwerse, 2012; Ozuru, Rowe, O'Reilly, & McNamara, 2008; van den

Broek, Lorch, & Thurlow, 1996) gradually develop throughout childhood. As such, a developmental perspective is essential in investigating individual differences in reading comprehension. Much less is known about the effects of working memory and comprehension processes on reading comprehension in older readers. Investigation of these effects in older readers is important because these children: (a) have already acquired decoding skills, and (b) fluently read connected text. It is possible that as readers get older working memory and ability to engage complex comprehension processes influence comprehension of increasingly difficult texts over and above vocabulary and background knowledge. Consequently, disentangling the exact nature of these relations is important as it might improve understanding of the developmental differences in reading comprehension as a function of reader characteristics, text features and specific comprehension processes tapped by a test item (i.e., text memory, text inference, knowledge integration). The design of the current study allows investigation of the effects of reader characteristics on comprehension within two developmental grade spans: (a) seven through nine and (b) ten through twelve. More specifically, capturing these effects in the context of developmental differences was possible because reader characteristics and reading comprehension measures were on the same metric within each grade span, but not across entire grade range from grade seven to grade twelve. The common metric was assured by administering the same testing battery to students in grades seven through nine and grades ten through twelve.

Text Features

Emerging literature suggests that comprehension skills not only depend on reader characteristics but are also a function of text features, and interaction of cognitive characteristics with these features (text-reader interaction; Linderholm, Everson, van den Broek, Mischinski, Crittenden, & Samuels, 2001). Text features entail properties of text including vocabulary load, linguistic structure, cohesion, discourse style, and genre. These features greatly affect comprehension as they determine the extent to which information included in the text is easily accessible to the reader. In general, poorly written and structured texts with many informational gaps are hard to understand. Additionally, texts can be difficult to understand if they draw on the reader's background knowledge as there might be a discrepancy between reader's depth and breadth of background knowledge and conceptual difficulty of the text. A detailed description of all text features is beyond the scope of the present study. We turn now to a discussion of selected text features including word frequency, sentence length, cohesion, and genre.

Word frequency, sentence length and text cohesion. Traditionally, research investigating effects of text difficulty on comprehension has predominantly focused on word frequency (signifying word difficulty) and sentence length, with less frequent (low frequency) words and longer sentences characterizing more difficult texts (Perfetti, 1985). This focus on word frequency and sentence length has been criticized because it largely ignores, except by proxy, inter-sentential factors that affect text difficulty. Not surprisingly, more recent research indicates that text difficulty is also a function of text cohesion which represents the extent to which ideas in the text are presented in an explicit and meaningful manner, with low cohesion texts being less explicit (with more conceptual and structural gaps) and more dependent on background knowledge for comprehension (Graesser, McNamara, Louwerse, & Cai, 2004; Graesser & McNamara, 2011; McNamara, Louwerse, McCarthy, & Graesser, 2010).

Cohesion is a function of various linguistic and discourse markers (explicit word, phrase, or sentence) which help readers in constructing meaningful ideas conveyed in the text (Graesser, McNamara, & Louwerse, 2003). Referential and deep cohesion are important sources of variation in text cohesion. High referential (i.e., the overlap/repetition of words and concepts across sentences, paragraphs, or an entire text) and deep cohesion (i.e., the amount of connecting words which are helpful in clarifying relations between events, ideas, and information) significantly improve understanding of the text as they decrease the necessity of using inferences to understand relations in the text (Graesser et al., 2003; McNamara, 2001). Low cohesion texts are more beneficial for readers with high background knowledge as they engage them in complex comprehension processes such as inference making and knowledge integration in order to fill in conceptual gaps. Low cohesion texts also stimulate deeper understanding of the text as well as better text recall in these readers. High cohesion texts are better for readers with low background knowledge as these texts are more explicit and less background knowledge dependent (McNamara & Kintsch, 1996). However, the explicit and straight forward nature of high cohesion texts reduces the need for making inferences and integrating textual information with background knowledge (McNamara, Kintsch, Songer, & Kintsch, 1996).

Genre. Text genre is typically divided into narrative, expository, and mixed, which simply combines characteristics of narrative and expository texts. The current study is restricted to narrative and expository texts. Narrative and expository texts not only differ in terms of text features including structure but also their purpose. Narrative texts are written to entertain readers as they have a gradually developing theme. Typically, this type of text is easier to understand as it draws on everyday events and experiences. Conceptual gaps can be filled in with greater ease because most readers (who are proficient at lower level comprehension processes) have relevant background knowledge to grasp ideas presented in the text. Consequently, simplicity of narrative texts is associated with use of familiar words (high frequency words) and connectives demarcating causal relations (McNamara, Graesser, & Louwerse, 2012). In contrast to narrative texts, the main purpose of expository texts is to educate readers about a specific topic. These texts are harder to understand because they convey potentially unfamiliar information unrelated to the typical reader's everyday experiences (McNamara et al., 1996). Consequently, the difficulty of expository texts is related to the use of unfamiliar words (low frequency words) and the inclusion of fewer connectives. Interestingly, these texts have high referential cohesion, which might be attributed to the repetitive presentation of taught concepts (McNamara et al., 2012). The different nature of narrative and expository texts requires that they be processed differently by the reader. Skillful readers process expository and narrative texts in distinct ways as expository texts are more dependent on background knowledge when compared with narrative texts (Linderholm & van den Broek, 2002; Vellutino, 2003).

Explanatory Item Response Models (IRT)

Typically, item response models (descriptive item response models) have been used to assess one's abilities, skills, or characteristics on the latent trait (construct) measured through a test. More recently, there has been a growing interest in using item response models to explain item responses on the test in terms of external variables associated with persons and/or items (de Boeck & Wilson, 2004). These models are distinct from descriptive item response models as they focus on relations between specific variables (test responses/test scores and external variables) rather than simply focusing on one's position on the latent trait. They are referred to as explanatory item response models because they attempt to explain one's responses to test items based on person and item parameters as well as external variables. Consequently, explanatory item response models attempt to simultaneously establish one's position on the ability dimension as a function of person characteristics, and an item's position on the difficulty dimension (i.e., the ability scale) as a function of item features. In other words, these models attempt to explain person's position on the ability dimension while also explaining why one item is more difficult than another, or more precisely, why an item obtains the specific difficulty value. Through this joint explanatory model, it is possible to specify where an unmeasured item will fall on the difficulty scale, and where an unmeasured person will fall on the ability scale.

Person parameters, item parameters and external variables. The person parameter (θ_p) represents an individual's location on the latent trait continuum, which is typically assumed to follow a normal probability distribution. The person parameter might also be interpreted in terms of individual differences between persons on that trait. In other words, the person parameter captures the person's ability on a scale, the characteristics of which are established by the psychometric model for the items. The item parameters ($\alpha_i \beta_i$) reflect the discriminability and difficulty of an item, respectively. In many applications, the discrimination indices are assumed equal across items, indicating that items differ only in their difficulty. This model is referred to as the Rasch model and is generally preferred over other item response models for ability tests because it results in a scale that is interval based and allows for the number correct score to serve as a sufficient statistic for person ability (i.e., it only matters how many items were answered correctly, not the specific pattern of items answered correctly). In item response models, item parameters are described as a general effect of the item and are independent of the ability of the person or persons responding to the item (Fischer & Molenaar, 1995). Depending on the model being applied in a given context, the effects of person and/or item parameters might be modeled as fixed or random. This distinction is really one of establishing the inference space over which conclusions about item and person parameters will apply. Specifically, are they restricted to the specific items and persons included in the study (i.e., fixed parameters), or do they apply across some universe of similar items and people of which those in the study are considered a representative sample (i.e., random parameters). The external variables represent person, item or person by item covariates which are used to explain individual differences regarding responses to test items through their influence on person ability and item difficulty (de Boeck & Wilson, 2004; van den Noortgate, de Boeck, & Meulders, 2003). These covariates can be understood as additional (to person and item parameters) factors affecting responses to test items. Depending on the model, the effects of external covariates might be fixed or random.

In the context of the current project test scores on items measuring reading comprehension depend not only on reading comprehension ability and the difficulty of items measuring comprehension, but might be also explained using: (a) person covariates including working memory, level of background knowledge, etc. (b) item covariates including type of text, type of comprehension process, etc., as well as (c) the interaction of person covariates and item characteristics, e.g., working memory with text type.

Item response models for binary data. Explanatory item response models are appropriate for binary, ordered-category, and nominal-category data. However, description of models for ordered-category and nominal-category data is beyond the scope of the present study. For binary data (where only two responses are possible: correct or incorrect), explanatory item response models can be modeled as logistic or normal-ogive models depending on whether the logit or probit link function is used (respectively) to link the expected value of the binary observation with the expected value of the underlying continuous variable (de Boeck & Wilson, 2004). These two approaches are, for all intents and purposes equivalent, and for that reason the normal-ogive model will not be discussed further. All modeling was carried out using the one-parameter logistic regression model (Rasch model), which assumes equal item discrimination, varying item difficulties across items, and models the log-odds of a correct response as a linear function of person and item covariates (see Appendix for more details).

The Rasch models are item response models which originated in 1960 as models for measurement (Fischer & Molenaar, 1995). In these models, the linear component representing a logarithmic transformation of the binary response (η_{pi}) is determined by the ratio of the person's ability and item difficulty ($\eta_{pi} = \theta_p/\beta_i$; de Boeck & Wilson, 2004). The probability of a correct response becomes higher with an increase in the person's ability or a decrease in the item's difficulty. Consequently, persons who are higher on the latent trait continuum have a higher probability of answering any given item correctly compared to persons who are lower on the latent trait continuum. Similarly, for individuals with the same ability, the probability of passing an item declines as the item difficulty increases. The Rasch model places item difficulty and person ability on the same scale. If the person parameter equals the item difficulty, then the probability that such a person will answer such an item correctly is equal to 0.5. The described relation between the probability of a correct response and the ability of the person answering the item is referred to as the item response function (a mathematical function reflecting the relation between one's position on the latent trait continuum and the probability of a correct response to an item measuring this trait (Reise, Ainsworth, & Haviland, 2005). The item response function has a characteristic "S" shape. The inflection point on the item response function represents the difficulty parameter of the item. The inflection point on a curve is that point where the slope of the curve, i.e., the first derivative of the curve, changes direction (von Davier & Carstensen, 2006). For item response functions, the slope of the function changes direction at that point where the probability of passing the item equals the probability of failing the item, i.e., at $p = 0.5$ (Reise et al., 2005).

Advantages of explanatory item response models. Explanatory item response models are advantageous when compared to standard psychometric models for tests or statistical models for ability because they: (a) take into account correlations among the items, (b) allow for estimation of individual differences, (c) allow for modeling of random and fixed effects of person and/or item parameters as they belong to a broader class of the generalized linear mixed models, (d) explain the probability of correct responses utilizing external variables, and (e) jointly model the probability of correct responses as a function of person and item characteristics. Joint modeling provides a more accurate estimation of individual differences in a latent trait as it simultaneously estimates effects of person and item properties (de Boeck & Wilson, 2004). That is, joint modeling estimates the probability of correct responses using all information in the data (person side or item side data). In the proposed study, application of these models is most suitable because it allows for the joint modeling of person and text characteristics and their interaction on individuals' comprehension of text.

The Present Study

Reading comprehension is a crucial ability permitting effective functioning in the society. Yet approximately 30% of school-aged children experience reading difficulties (National Assessment of Educational Progress [NAEP], 2013). These problems result from deficient deployment of comprehension processes, and a resulting inability to construct a coherent mental representation of the text. Prior research has shown that reading comprehension depends on comprehension processes, reader characteristics as well as lexical and syntactic features of the text (Lorch & van den Broek, 1997; van den Broek & Espin, 2012). Much less is known about interactive effects of comprehension processes, reader characteristics and text features on individual differences in reading comprehension. The majority of previous findings rely on group data which limit the study of within-person variability in reading comprehension. The explanatory IRT allows investigating effects of text and reader characteristics from an individual differences perspective. The current paper used explanatory IRT models to investigate effects of reader characteristics, text features, comprehension processes, and the extent to which these interact in their effects on reading comprehension in 11 to 20 year old students. These models were employed to explain variation in test scores of reading comprehension items as a function of person and item characteristics as well as their interactions. Application of explanatory item response models to an operational test (Gates-MacGinitie Reading Test) in order to investigate individual differences in reading comprehension was novel, and allowed us to examine how students with different levels of various abilities perform on the operational test with correlated text features.

We expected to show that reader characteristics would be the most important determinants of test scores as good readers have a wider range of skills and consequently perform better on reading comprehension tests regardless of the text they read. That is to say, we predicted that accurate and fluent word reading, working memory, vocabulary and background knowledge would be the most influential in explaining reading comprehension. We further hypothesized that text features would explain some of the variation in item difficulties, but item difficulties would vary less than person abilities in the two developmental grade spans. Most importantly, we predicted that the interactions of text and reader characteristics would explain some of the within-person variability in test items, but for the most part text characteristics would not differentially affect item difficulties for readers with different attributes. That is, we expected that good readers would read well across the set of texts and items used in this study, and that difficult items were difficult for all readers in similar ways.

Method

Participants

The participants were evaluated in a larger study examining cognitive processes related to reading comprehension. Students were recruited from eight schools within four districts. Table 1 presents demographic information for the total sample and two blocks of students (grades 7-9 and grades 10-12). The total sample size included 1,190 students aged 11 to 20 years old (M = 14.97; SD = 1.73). The percentage of boys (50.92%) and girls (49.08%) was similar. The percentage of students within seventh (15.13%), eighth (16.39%), ninth (17.14%), tenth (19.16%), eleventh (17.90%), and twelfth (14.29%) grades was comparable. Students in grades 7-9 and 10-12 were comparable in terms of gender, ethnicity and qualifying for free meals.

Table 1. Demographic information for the total sample and by two blocks of students

	Grades 7-9 *n* = 579	Grades 10-12 *n* = 611	Total *n* = 1190
Age, *M (SD)*	13.65 (1.04)	16.44 (1.01)	14.97 (1.73)
Gender, % male	52.85	49.10	50.92
Ethnicity			
% Hispanic	54.40	47.46	50.84
% African American	23.14	20.95	22.02
% White	18.65	27.66	23.28
% Asian or Pacific Islander	1.55	2.95	2.27
% Other	2.27	0.98	1.59
Free meals, % qualified	58.68	53.03	55.77

Measures

Test of Word Reading Efficiency. The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner & Rashotte, 1999) was comprised of two subtests: sight word efficiency and phonemic decoding efficiency. In the current study, form A of each subtest was individually administered to students. The sight word efficiency subtest measured accuracy and speed of reading real words by instructing students to read out loud list of 104 increasingly difficult words. The phonemic decoding efficiency subtest (including 63 items) was used to assess accuracy and fluency of pronouncing phonemically regular non-words. For both subtests, the number of words read correctly within 45 seconds was recorded. The reliability of the subtests was 0.91 to 0.97, respectively. In the current project, a composite score representing a measure of reading fluency was computed by averaging the number of words read correctly per second on sight word efficiency and phonemic decoding efficiency subtests. This score was calculated because: (a) the correlation between the two TOWRE subtests was moderate ($r(1,235) = 0.64$, $p < 0.001$), and (b) the interest of the current project was on overall effects of fluency on comprehension, rather than on differential effects of real- and pseudo-word reading..

Woodcock & Johnson Tests of Cognitive Abilities - Letter Word Identification. The Letter Word Identification subtest of the Woodcock & Johnson Tests of Cognitive Abilities battery (Woodcock, McGrew, & Mather, 2007) was individually administered to measure decoding accuracy of real words. Students were asked to name letters and then read aloud words from a list of increasingly difficult words (words appearing less frequently in written English). The total number of correctly read words was recorded. Within sample Cronbach's alphas for the subtest for grades 7 to 12 ranged from 0.89 to 0.92.

Woodcock & Johnson Tests of Cognitive Abilities - Numbers Reversed. The Numbers Reversed subset of Woodcock & Johnson Tests of Cognitive Abilities battery (Woodcock, McGrew, & Mather, 2007) was individually administered to measure working memory. The task consisted of seven levels including trials with strings of digits of variable length. The first two levels (5 trials each) were comprised of strings of 2 and 3 digits, whereas levels 3 through 7 (4 trials each) were comprised of strings of 4 to 8 digits, respectively. Each trial started with the oral presentation of a string of digits to a student. Immediately after presentation, the student was required to recall and to repeat the string of digits in reverse order. The strings of digits were read to a student until the student responded incorrectly to three consecutive highest-numbered trials in a level. The total number of correct trials indicated was recoded. Within sample Cronbach's alphas for the subtest for grades 7 to 12 ranged from 0.75 to 0.83.

Gates-MacGinitie Reading Tests (GMRT) – Vocabulary Subtest. The Gates-MacGinitie vocabulary subtest (MacGinitie, MacGinitie, Maria, & Dryer, 2000) was group administered to measure student's preexisting knowledge of common and rare words. Form S, grade levels 7/9 and 10/12 were used in the current study. Students were presented with words placed within a context suggesting a part of speech of the word, and then were asked to select the most appropriate word or phrase related to the target word. The total number of items for each form was equal to 45. The total number of correctly answered items was recoded. Within sample Cronbach's alphas for the subtest for grades 7 to 12 ranged from 0.83 to 0.90.

Gates-MacGinitie Reading Tests – Background Knowledge (GMBK). The group administered GMRT background knowledge test (Barnes & Watkins, 2013) was used to measure students preexisting word and world knowledge considered necessary to comprehend the GMRT reading comprehension passages. Two forms corresponding to the GMRT reading comprehension subtest for grade levels 7/9 (form S) and 10/12 (form S) were created. Construction of each form involved three-stages: (a) reading of the GMRT reading comprehension passages, (b) establishing vocabulary and background knowledge necessary to answer reading comprehension questions, and (c) evaluating whether word and world knowledge questions mapped to background knowledge considered as helpful in comprehending passages and answering corresponding to them questions. The total number of items for grade level 7/9 and 10/12 was equal to 29 and 35, respectively. The total number of correctly answered word and world knowledge items was recorded. Within-grade Cronbach's alphas (for grades 7 to 12) ranged from 0.55 to 0.63 for raw scores on the form for grade levels 7/9, and ranged from 0.58 to 0.77 for raw scores on the form for grade levels 10/12.

In the current project, a single background knowledge score was constructed by adding correct responses to word and world knowledge items in subtests for grade levels 7/9 and 10/12, and dividing by the total number of items in these two subtests (i.e., 64 items). This percent correct score across all GMBK items was computed because results of the factor analysis involving word and world knowledge scores for levels 7/9 and 10/12 indicated presence of one general factor. Only the first factor had an eigenvalue greater than one (i.e., eigenvalues of the first factor for grades 7 through 9 and grades 10 through 12 were equal to 1.52 and 1.85, respectively).

Gates-MacGinitie Reading Tests (GMRT) – Reading Comprehension Subtest. The GMRT reading comprehension subtest (MacGinitie, MacGinitie, Maria, & Dryer, 2000) was group administered to measure student's ability to read and comprehend texts differing in both length and subject matter. Form S, grade levels 7/9 and 10/12 were used in the current study. Students were asked to a read passage and then answer multiple choice questions related to the information presented in the passage. Providing correct answers to questions was dependent on ability to recall information provided in the passage, and ability to draw inferences. The total number of passages for each form was equal to 11, while the total number of questions for each form was equal to 48. Within sample Cronbach's alphas for the subtest for grades 7 to 12 ranged from 0.89 to 0.94.

Coding

Coding of items. Ninety-six individual questions from the GMRT reading comprehension subtest for grade levels 7/9 (form S) and 10/12 (form S) were coded in terms of comprehension processes employed by readers to provide correct answers to questions. Three types of comprehension processes were distinguished. The first type comprised *text memory questions*, for which a correct answer was obtained by literal recall of information from the text. Providing correct answers to questions did not require background knowledge (beyond the meaning of the word in the sentence) and did not necessitate extrapolating or combining information presented across different sentences within a passage other than through recall. Of note, a fact presented in the passage that was later negated or reworded in the question was considered as a memory item. The second item type involved *text-based inference questions*. Answering this type of question required making novel inferences based on information provided in different sentences within a passage, and integrating information from the text (if necessary from a whole passage). No background knowledge was necessary to answer questions correctly. The third type included *the knowledge-based inference questions* for which providing correct answers required integrating information presented in the text with the reader's background knowledge.

This coding scheme, as opposed to other coding schemes – for instance, Anderson (1972), was applied for two reasons. First, it was constructed on the basis of one of the most notable reading comprehension models focusing on comprehension processes, namely the construction-integration model (Kintsch, 1988). This model presumes that text memory and text based-inferences are deeper comprehension processes involved in creating the text-based representation of the text, whereas knowledge-based inferences are deeper comprehension processes necessary to construct the situation model of the text integrating text-based and background knowledge information. Second, the proposed coding scheme was based on the work of Hannon and Daneman (2001) which has proven to be successful in measuring comprehension processes in prior experimental work with college students.

Two coders were provided with instructions and a coding key to categorize questions from the GMRT reading comprehension subtest for grade levels 7/9 (form S) and 10/12 (form S). During the first phase of the coding process, the coders separately coded all of the questions based on the provided materials. The raters agreed on 75% of all questions for grade level 7/9, and 81% of all questions for grade level 10/12. The weighted kappa for grade level 7/9 was equal to .51, and the weighted kappa for grade level 10/12 was equal to .58. During the second phase of the coding process, the raters and the mediator met in order to discuss discrepancies in coding and reached agreement regarding discrepant items.

Coding of passages. Twenty two passages from the Gates MacGinitie reading comprehension subtest for grade levels 7/9 (form S) and 10/12 (form S) were coded as narrative or expository. The text was considered as narrative if it included a gradually developing theme and drew on everyday events and experiences. The text was classified as expository if its nature was scientific and conveyed unfamiliar information unrelated to everyday experiences.

Measurement of passage features. The current project focused on measuring the following text features: word frequency, sentence length as well as referential and deep cohesion. The average word frequency, average sentence length, and overall text difficulty of the GMRT reading comprehension passages were measured utilizing the *Lexile Text Analyzer* software (Schnick & Knicklebine, 2007). The average log word frequency measure as opposed untransformed average word frequency was used because the logarithmic transformation corrects the distribution of word frequency so that it approximates a normal distribution, as well as has a linear fit with reading times (Graesser et al., 2004). The Lexile measure estimated the overall text difficulty based on word frequency and sentence length. This measure is a numeric representation of complexity of the text and reader ability, and is largely a function of sentence length and word frequency. Cohesion characteristics of the GMRT reading comprehension passages were measured utilizing the *Coh-Metrix Text Easability Assessor* (a computational tool providing five text easibility scores; Graesser, McNamara, & Kulikowich, 2011). Referential cohesion was measured with the referential cohesion score expressed as a percentile whereas deep cohesion was assessed using the deep cohesion score expressed as a percentile.

Data Analyses

All hypotheses were evaluated with explanatory, dichotomous, multivariate, cross-classified extension of the Rasch model with random intercepts. The dichotomous form of the Rasch model was applied because item responses were in a correct-incorrect format (missing values where coded as incorrect responses). A multivariate framework was used as the item difficulties of all reading comprehension items were simultaneously modeled. A cross-classification structure was used to deal with dependencies among the responses to items, as these dependencies resulted from administering all items to all students with students responding to all items. Thus, item responses were cross-classified in persons and items (i.e., there are 2 levels of nesting with two crossed factors at level 2: item responses at level 1 are nested simultaneously within items and persons at level two). In the current modeling, the nesting of items within passages is ignored. The proposed models can be described as cross-classified linear logistic test models with: (a) 1st-level model including responses to items (dummy variables where 1=correct, 0=incorrect), (b) 2nd-level including item and person parameters which are crossed in their design. The Maximum likelihood estimation based on Laplace approximation was used to estimate unknown parameters. All models were estimated using the *glmer* function of *lme4* package in *R* (Bates, Maechler, & Dai, 2008).

To be clear, the analytic models employed in the study were models that aimed to explain individuals' responses to comprehension test items. The models indicated that the probability of a particular person correctly answering a particular test item was a function of two things: the difficulty of the item and the ability of the individual. The model explicated the features of test questions and text passages that determined item difficulty, and the characteristics of readers that determined their ability to comprehend text.

Table 2 present person characteristics, text features and text-reader interactions used in the different classes of estimated models. In all models, person and item parameters were random, whereas effects of person and item covariates were fixed. The models were separately computed for two blocks of students. The first block comprised students from grades seven through nine, and the second block included students from grades ten through twelve. Assigning students to two blocks was done to ensure that person covariates are on the same metric. Because there are no common items across the two-forms of the GMRT, it is not possible to jointly model the two forms of the test and place all item parameters on a common scale. Thus, modeling each form separately is preferred to joint modeling. A stepwise modeling approach was used to address hypotheses (see Table 2). The fully unconditional model without person and item covariates was used to estimate residual variance on person and item side of the data. After that, the models for reader characteristics (model 1) and passage features (model 2) were estimated to examine the effects of reader characteristics and passages features on difficulty of reading comprehension items, respectively. The last model (model 3) estimated interactive effects of text and reader characteristics. Interactive effects included in the third model were not exhaustive as many more could have been entered. However, they were a good representation of possible text-reader interactions based on previous findings in the literature. In all models, continuous reader and text predictors were grand mean centered within two grade blocks (grades 7- 9 and grades 10-12) in order to provide a correct and meaningful interpretation of estimates.

Table 2. Reader characteristics, text features and text-reader interactions in estimated models

READER CHARACTERISTICS (Model 1)			
Construct	**Measure**	**Subtest**	**Test score**
Word reading	WJIII	Letter word identification	Raw score
Reading fluency	TOWRE	Composite score	Raw score
Vocabulary	GMRT	Vocabulary	Raw score
Background knowledge	GMBK	Composite score	Raw score
Working memory	WJIII	Numbers reversed	Raw score
PASSAGE FEATURES (Model 2)			
Construct	**Calculation**	**Test score**	
Word frequency	Coh-Metrix version 3.0	Mean log frequency	
Sentence length	Coh-Metrix version 3.0	Mean sentence length	
Genre	Examiner's coding	Categorical (expository or narrative)	
Referential cohesion	Coh-Metrix Text Easability Assessor	Percentile	
Deep cohesion	Coh-Metrix Text Easability Assessor	Percentile	
TEXT-READER INTERACTIONS (Model 3)			
Reader Characteristic	**Item Type**	**Passage Feature**	
Vocabulary	Comprehension Process	None	
Background knowledge	Comprehension Process	None	
Working memory	Comprehension Process	None	
Vocabulary	None	Referential cohesion	
Background knowledge	None	Referential cohesion	
Working memory	None	Referential cohesion	
Vocabulary	None	Deep cohesion	
Background knowledge	None	Deep cohesion	
Working memory	None	Deep cohesion	
Vocabulary	None	Word frequency	
Background knowledge	None	Word frequency	
Working Memory	None	Sentence length	
None	Comprehension Process	Referential cohesion	
None	Comprehension Process	Deep cohesion	

Results

Table 3 provides descriptive statistics for person characteristics by two grade blocks (grades 7-9 and grades 10-12). Not surprisingly given the sample sizes and large age span differences between the two blocks of students, statistically significant differences between students in these blocks were observed in word reading ($t(1,188) = -17.81$, $p < 0.001$), reading fluency ($t(1,105) = -10.27$, $p < 0.001$), vocabulary ($t(1,101) = -4.35$, $p < 0.001$), background knowledge ($t(1,163) = -16.43$, $p < 0.001$), and working memory ($t(1,110) = -4.83$, $p < 0.001$). These findings suggested that students in grades 10 through 12 had higher reading ability, vocabulary and background knowledge, as well as working memory capacity when compared with students in grades 7-9.

Table 8 (in the appendix B) presents bivariate correlations between reader characteristics in two developmental grade spans. For both grades spans, moderate to large positive correlations were found between word reading and reading fluency, vocabulary, background knowledge, and working memory capacity (all between 0.43 and 0.54 across both grade spans). Reading fluency was weakly correlated with vocabulary (0.26), background knowledge (0.24) and working memory (0.25) in grades 7 through 9, and moderately correlated with these measures in grades 10 through 12 (0.31, -0.33). Background knowledge was highly correlated with vocabulary (0.70, 0.76) and moderately correlated with working memory (0.28, 0.22) in both grade spans.

Table 3. Descriptive statistics for reader characteristics in two grade-span blocks

	Grades 7-9			Grades 10-12			Cohen's d
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	
Word reading	579	61.56	4.43	611	65.58	3.31	1.03
Reading fluency	561	1.36	0.20	546	1.48	0.21	0.59
Vocabulary	559	21.22	7.56	544	23.30	8.31	0.26
Background knowledge	568	0.62	0.11	597	0.73	0.11	1.00
Working memory	564	12.92	3.17	548	13.86	3.31	0.29

Table 4 contains descriptive statistics for item and passage characteristics by two grade levels of the GMRT reading comprehension test (grade level 7-9 and grade level 10-12). There were statistically significant differences between the two levels of GMRT passages in average log word frequency ($t(94) = -2.32$, $p < 0.05$), Lexile measure ($t(94) = 2.13$, $p < 0.05$), and referential cohesion ($t(94) = 2.04$, $p < 0.05$). Passages for grades 10 through 12 were more difficult as measured by the Lexile scale, and also had higher referential cohesion as well as lower average log word frequency. The findings indicated that these passages were harder to understand as they used less familiar words and had a higher level of text difficulty. At the same time, their higher referential cohesion suggested greater overlap of explicit words and ideas in text which was helpful in understanding more challenging content. The two levels of GMRT passages were not statistically different in terms of the average sentence length ($t(94) = 1.01$, p = 0.32), genre ($\chi^2(1) = 0.68$, p = 0.84), deep cohesion ($t(94) = -0.70$, p = 0.49), and specific comprehension processes tapped by test items ($\chi^2(1) = 0.04$, p = 0.06). Of note, the average sentence length and Lexile measure were highly correlated ($r(96) = 0.89$, p < 0.001) introducing a problem of multicollinearity. To alleviate this problem the Lexile measure was not used as a predictor in the estimated models.

Table 4. Descriptive statistics of item and passage features

	Form 7-9 *n of questions* = 48	Form 10-12 *N of questions* = 48
Word frequency, *M* (*SD*)	3.56 (0.20)	3.45 (0.24)
Sentence length, *M* (*SD*)	19.73 (5.32)	20.86 (5.50)
Lexile measure, *M (SD)*	1099.79 (174.08)	1183.33 (204.47)
Referential cohesion, *M* (*SD*)	27.04 (23.27)	39.04 (32.92)
Deep cohesion, *M* (*SD*)	55.25 (28.29)	50.54 (29.17)
Item type/Comprehension process		
% Text memory	50	70.83
% Text inference	50	29.17
Genre		
% Narrative	43.75	47.92
% Expository	56.25	52.08

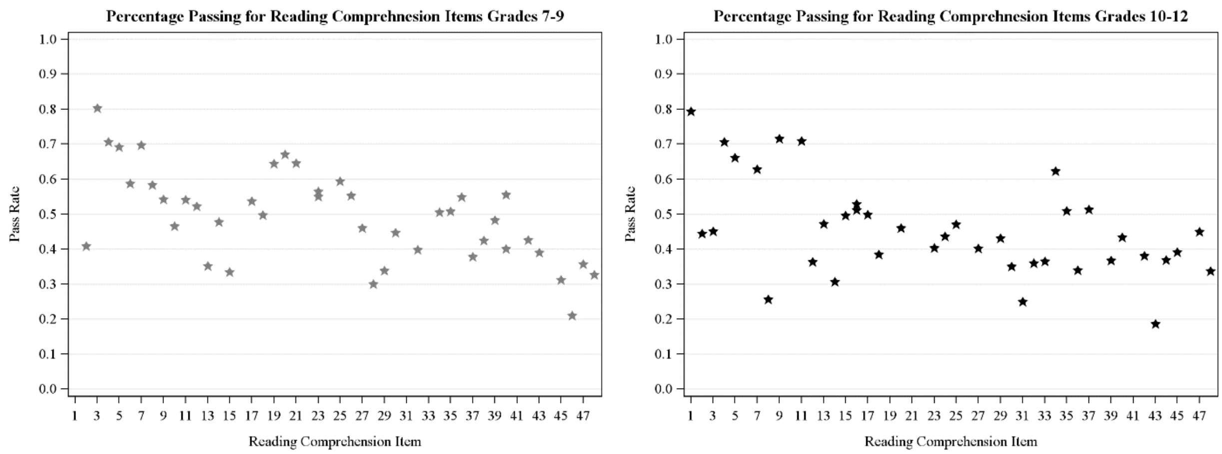
Table 5 presents estimates of residual variance on the person and item sides of the data for estimated models. The variance components models (i.e., fully unconditional model) were used to estimate the unexplained variation in reading comprehension item scores as a function of person abilities and item difficulties. For both grade blocks (grades 7-9 and grades 10-12), the results suggested that the unexplained variance was higher for person abilities than item difficulties. That is to say, there was more variance to be explained in person abilities than in item difficulties. This finding was not surprising, as item difficulties generally vary less than person abilities by design. If items were allowed to vary beyond the range of person abilities, there would be an accumulation of items that all persons answered correctly and items that no one answered correctly. Estimated difficulties of these items would be biased toward the mean item difficulty, and the variance in item difficulties would be underestimated. More importantly, these items would provide limited information about person ability.

Adding reader characteristics to the models (i.e., model 1) substantially decreased the unexplained person variance but not the unexplained item variance (relative to the fully unconditional models). These results were expected as reader characteristics explain person abilities and item features explain item difficulties. At the same time, adding passages characteristics to the models (i.e., model 2) decreased the residual item variance but not the residual person variance (relative to the initial models). Again, this finding was expected as these models were designed to explain the residual variance on the item side of the data, but not on the person side of the data. Together, item characteristics and person characteristics explain item responses and the probability of a particular person answering a particular item correctly. Simultaneous inclusion of reader characteristics, item and passage characteristics, as well as their interactive effects in the models (i.e., model 3) decreased the unexplained item variance (relative to the initial model and model 2) and person variance (relative to the initial model and model 1), with a larger decrease in the residual item variance. These results suggested that interactive effects were more important for explaining item difficulties than person abilities. That is to say, the interaction of person and item characteristics indicates that the impact of item characteristics on item difficulties were not uniform across all readers, or alternatively, that the impact of person characteristics on reading comprehension were not uniform across all item and text types.

Table 5. Residual variance on the person and item side of data in estimated models

	GRADES 7-9					
	Person Side			**Item side**		
Estimated Model	*Variance*	*SE*	*Variance reduction*	*Variance*	*SE*	*Variance reduction*
Fully Unconditional	1.42	0.10	NA	0.48	0.10	NA
Model 1 - Reader Characteristics	0.46	0.04	0.68	0.48	0.10	0.00
Model 2 - Passage Features	1.42	0.10	0.00	0.37	0.08	0.23
Model 3 - Text-Reader Interactions	0.43	0.03	0.70	0.26	0.04	0.46
	GRADES 10-12					
	Person Side			**Item side**		
Estimated Model	*Variance*	*SE*	*Variance reduction*	*Variance*	*SE*	*Variance reduction*
Fully Unconditional	3.28	0.24	NA	0.71	0.15	NA
Model 1 - Reader Characteristics	0.50	0.04	0.85	0.71	0.15	0.00
Model 2 - Passage Features	3.29	0.24	0.00	0.51	0.11	0.28
Model 3 - Text-Reader Interactions	0.49	0.04	0.85	0.40	0.07	0.44

Figures 1 and 2 present the pass rates for GMRT reading comprehension items in grades 7-9 and grades 10-12, respectively. The pass rates were estimated based on frequencies of correct responses to these items. For both grade blocks, the pass rates ranged from 0.20 to 0.80, with more difficult items appearing towards the end of the test. The variation in pass rates was larger for grades 10-12 than for grades 7-9. This finding may reflect a greater range in abilities among the older students, or wider variation in the sampling of item difficulties in developing the test for older readers. The results in Table 5 suggest that there was greater unexplained person and item variance in the older grades (see results for the initial models in table 5). That is to say, there was greater variation in both reading comprehension ability among readers and greater variability in test difficulty in tests items for grades 10-12. However, because the ability scales and difficulty scales are not constrained equal across the two grade blocks, these variance estimates are not directly comparable across the two grade blocks.



Figures 1-2. Dot plots showing actual pass rates for the GMRT reading comprehension items for grades 7-9 and 10-12. Stars were used to represent the pass rate for a given item.

The initial findings were further explored utilizing the main effects models and models with interactive effects for grades 7-9 and grades 10-12. The main effects models attempted to explain variation in reading comprehension test items as a function of person abilities (i.e., model 1) or item difficulties (i.e., model 2). The models with interactive effects (i.e., model 3) were used to explain variability in reading comprehension test items through joint modeling of person's position on the ability dimension and item's position on the difficulty dimension.

Table 6 presents estimates of effects of person characteristics on the log odds of answering an average item correctly and the estimated pass rates for individuals of *low, average,* and *high* ability on a particular reader characteristic, specifically word reading, decoding fluency, vocabulary, background knowledge, and working memory. The pass rates on the GMRT reading comprehension test for students with low, average and high ability were derived by transforming the estimated log odds into a probability of passing an item of average difficulty. To obtain the estimated pass rate, ability scores for individuals with low and high ability were computed by first subtracting or adding (respectively) one standard deviation to an average ability score on that person characteristic in that grade span. Afterwards, the pass rates for readers who were at different positions on the latent ability continua were calculated by: (1) multiplying the estimated effect of the person characteristic on the log odds (i.e., the logistic regression parameter in Table 6) and the previously computed ability scores for individuals of low, average, and high ability. The resultant product gave the expected adjusted log odds for an individual of low, average, or high ability. The pass rate was obtained by transforming the expected adjusted log odds to a probability by computing $\frac{\exp(\text{expected adjusted log odds})}{1+\exp(\text{expected adjusted log odds})}$.

The model for grades 7-9 indicated that vocabulary, background knowledge and working memory were statistically significant predictors of reading comprehension scores over and above other predictors. Word reading and reading fluency were not statistically significant, conditional on the other person characteristics in the model. The pass rates for students with high vocabulary, background knowledge and working memory were higher when compared to the pass rates for students who were low on these ability continua. At the same time, the model for grades 10-12 indicated that reading fluency, vocabulary, and background knowledge were the most important for predicting reading comprehension scores, holding other predictors constant. Working memory and word reading were not statistically significant. The pass rates for students with high reading fluency, vocabulary, and background knowledge were higher than the pass rates for students who were low on these ability continua. That is to say, the obtained findings suggested that readers with stronger skills were better at comprehending texts, with strong semantic skills being the most important determinants regardless of grade level.

Table 6. Estimated effects of person characteristics on the log odds and estimated pass rates for low, average, and high ability readers in models including only person characteristics.

Measures	Log odds	Std. Error	Pass Rate Low Ability	Pass Rate Average Ability	Pass Rate High Ability
Grades 7-9 (N = 545)					
Word Reading	0.01	0.01	0.49	0.50	0.51
Reading fluency	0.31	0.19	0.48	0.50	0.52
Vocabulary	0.07***	0.01	0.37	0.50	0.63
Background knowledge	1.79***	0.41	0.45	0.50	0.55
Working memory	0.03*	0.01	0.48	0.50	0.52
Grades 10-12 (N = 531)					
Word Reading	0.02	0.01	0.49	0.50	0.51
Reading fluency	0.58**	0.19	0.47	0.50	0.53
Vocabulary	0.07***	0.01	0.35	0.50	0.65
Background knowledge	1.38**	0.48	0.46	0.50	0.54
Working memory	0.02	0.01	0.49	0.50	0.51

Note: $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

Table 7 contains log odds and pass rates for passage characteristics estimated in the models including word frequency, sentence length, genre, referential and deep cohesion measures. The pass rates on GMRT reading comprehension test for passages with *low* and *high* word frequency, sentence length, referential and deep cohesion were derived by transforming the estimated log odds into a probability. The steps involved in this transformation were analogous to the steps described for models including reader characteristics. The models for grades 7-9 and 10-12 indicated that narrative and expository passages differed in item difficulties, over and above the effects of other predictors. Expository passages were more difficult than narrative passages, and had a pass rate below 0.50. Additionally, the model for grades 10-12 suggested that deep cohesion was a statistically significant predictor of item difficulty holding other predictors constant. Items from passages that were low in deep cohesion had lower pass rates when compared to items from passages with high deep cohesion. This finding suggested that a greater amount of connective words clarifying relationships between text information significantly improved passage comprehension.

Table 7. Estimated effects of passage characteristics on the log odds and estimated pass rates for low and high item difficulties in models including only passage characteristics.

Grades 7-9 (N = 578)				
Measures	Log of Odds	Std. Error	Pass Rate Low	Pass Rate High
Word frequency	-0.71	0.94	0.54	0.47
Sentence length	-0.05	0.04	0.57	0.43
Expository Passage[a]	-0.52*	0.21	0.37	
Narrative Passage[a]	0.41	0.25	0.60	
Referential cohesion	0.01	0.01	0.44	0.56
Deep cohesion	0.01	0.01	0.43	0.57
Grades 10-12 (N = 611)				
Measures	Log of Odds	Std. Error	Pass Rate Low	Pass Rate High
Word frequency	-0.68	0.62	0.54	0.46
Sentence length	0.01	0.03	0.49	0.51
Expository Passage[a]	-0.51*	0.18	0.38	
Narrative Passage[a]	-0.35	0.19	0.41	
Referential cohesion	-0.01	0.04	0.58	0.42
Deep cohesion	0.01*	0.01	0.43	0.57

Note: *p < .05 **p < .01 ***p < .001
[a]Expository and narrative passage variables were dichotomous

The last set of models (model 3) was used to explain variation in reading comprehension test scores as a function of text-reader interactions. These models were used to determine whether within-person variability in performance on reading comprehension items might be further explained by interactive effects of text and reader characteristics over and above their respective main effects. The model for grades 7-9, revealed statistically significant main effects of genre (b = -0.76, SE = 0.35, p < 0.05) and background knowledge (b = 1.67, SE = 0.44, p < 0.001) holding other predictors and interactions constant. Expository passages were more difficult than narrative passages as the probability of correct responses to items associated with these passages was 0.32 lower. The pass rates for students with high background knowledge ($\pi = 0.55$) were higher when compared with pass rates for students with average ($\pi = 0.50$) or low background knowledge ($\pi = 0.45$). The background knowledge deemed to be necessary to understand relevant GMRT reading comprehension passages was an important determinant of reading comprehension test scores.

Figures 3 through 6 provide graphical depictions of statistically significant text-reader interactions for grades 7 through 9. There were statistically significant interactions between: (1) referential cohesion and comprehension processes tapped by a test item (b = -0.02, SE = 0.01, $p < 0.01$), (2) referential cohesion and vocabulary (b = -0.0005, SE = 0.0001, $p < 0.001$), (3) word frequency and vocabulary (b = -0.06, SE = 0.02, $p < 0.001$), and (4) deep cohesion and working memory (b = -0.0003, SE = 0.0001, $p < 0.05$).

The interaction of referential cohesion with comprehension processes tapped by a test item suggested that the probability of correct responses to text memory and text inference items varied depending on the level of referential cohesion. However, the interactive effects were dominated by the main effects. As such, the pass rates for text memory and text inference items were lower for low referential cohesion passages when compared with high referential cohesion passages. In other words, low referential cohesion passages were more difficult regardless of the specific comprehension process (i.e., text memory or text inference) tapped by the test item. Interestingly, text memory items were more difficult than text inference items regardless of the level of referential cohesion in the passage.

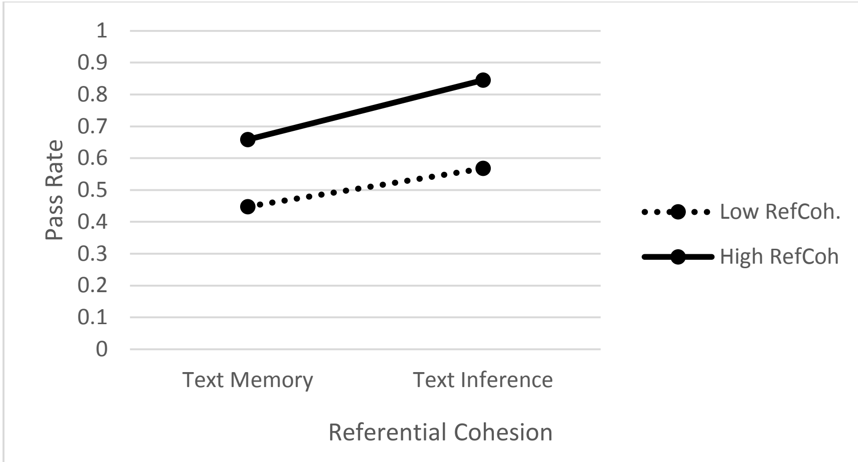


Figure 3. A line plot demonstrating pass rates for text memory and text inference items for low and high referential cohesion passages. Low RefCoh. = low referential cohesion; High RefCoh. = high referential cohesion.

The interaction of referential cohesion and vocabulary suggested that the effects of referential cohesion varied depending on students' vocabulary. However, the interactive effects were small when compared with the predominant main effects. As such, students with low vocabulary had lower pass rates than students with average or high vocabulary on passages with low or high referential cohesion. That is to say, readers with low vocabulary performed worse than readers with average or high vocabulary regardless of level of referential cohesion. Similarly, all students performed better on passages with high referential cohesion, although the effects of referential cohesion were somewhat more pronounced for students with low vocabulary, i.e., high vocabulary students showed less of a difference between high and low referential cohesion passages.

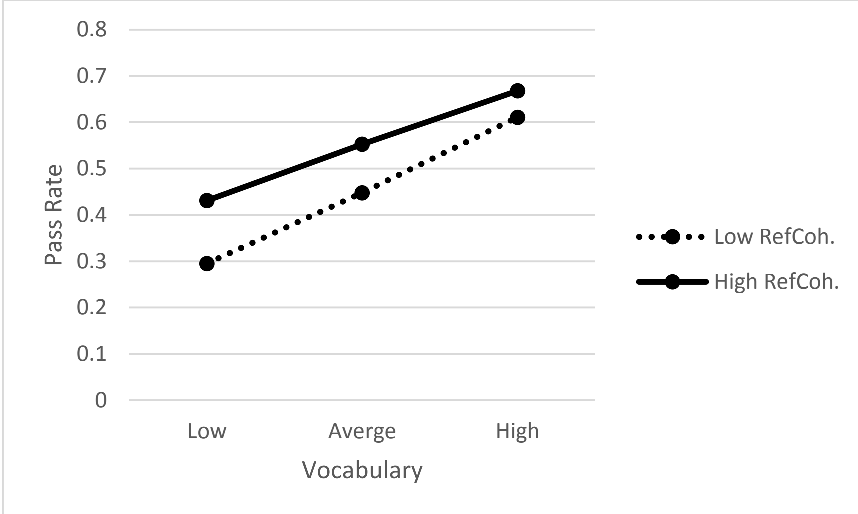


Figure 4. A line plot demonstrating pass rates for low and high referential cohesion passages among students with low, average and high vocabulary skills. Low RefCoh. = low referential cohesion; High RefCoh. = high referential cohesion.

The interaction of word frequency and vocabulary revealed that the effects of word frequency varied depending on students' vocabulary. However, the interactive effects were dominated by the main effects. Specifically, students with low levels of vocabulary had lower pass rates than students with average or high levels of vocabulary on passages with low or high word frequency. Consequently, readers with low vocabulary performed worse than readers with average or high vocabulary regardless of the level of text difficulty defined in terms of average word frequency. Again, the effects of word frequency on comprehension were mitigated by vocabulary, such that the effects of word frequency were negligible for high vocabulary students and more pronounced for students with weaker vocabularies.

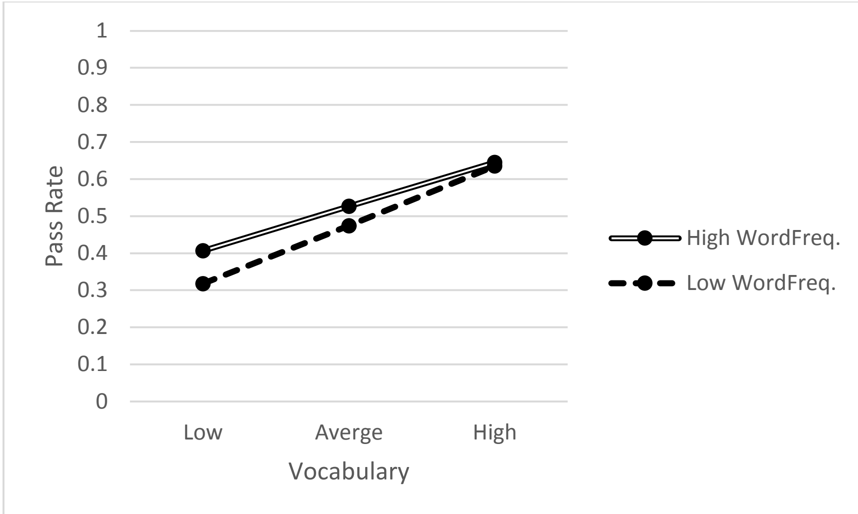


Figure 5. A line plot demonstrating pass rates for low and high word frequency passages among students with low, average and high vocabulary. Low WordFreq. = low word frequency; High WordFreq. = high word frequency.

Finally, the interaction of working memory with deep cohesion suggested that the effects of deep cohesion varied depending on students' working memory. However, the interactive effects were small when compared with the predominant main effects. As such, the findings suggested that both low and high deep cohesion passages were easier for students with high working memory capacity as compared to students with low working memory capacity. In other words, the results revealed that students with high working memory capacity had a higher chance of providing correct responses to reading comprehension test items regardless of the level of deep cohesion in the passage on which the test item was based.

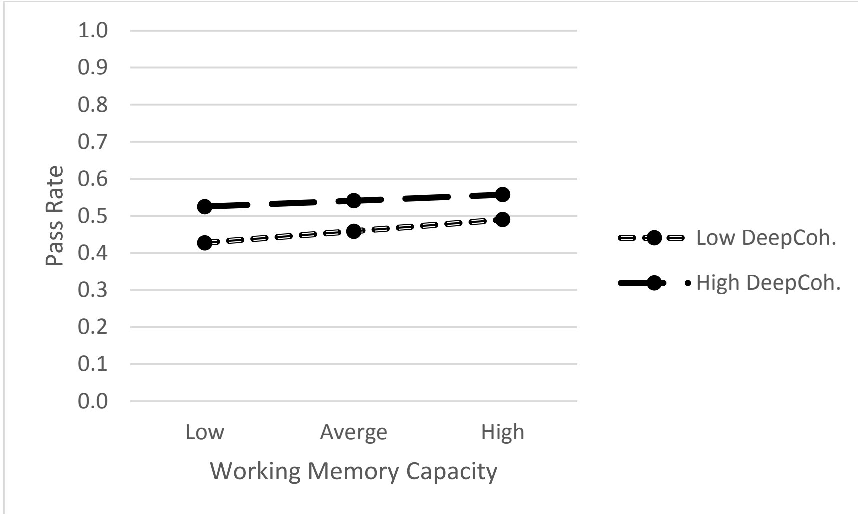


Figure 6. A line plot demonstrating pass rates for low and high deep cohesion passages among students with low, average and high working memory capacity. Low DeepCoh. = low deep cohesion; High DeepCoh. = high deep cohesion.

The model for grades 10-12 indicated statistically significant main effects of reading fluency (b = 0.50, SE = 0.19, p < 0.05) and vocabulary (b = 0.07, SE = 0.01, p < 0.001) holding other predictors and interactions constant. The pass rates for students who read more fluently ($\pi = 0.53$) and had higher vocabulary ($\pi = 0.65$) were higher when compared to pass rates for students who had an average level of fluency ($\pi = 0.50$) and vocabulary ($\pi = 0.50$), or students whose fluency ($\pi = 0.47$) and vocabulary ($\pi = 0.35$) skills were low.

Figures 7 and 8 provide graphical depictions of statistically significant text-reader interactions for grades 10 through 12. There were statistically significant interactions between: (1) working memory and the comprehension processes tapped by a test item (b = 0.03, SE = 0.01, p < 0.01), and (2) referential cohesion and background knowledge (b = -0.02, SE = 0.01, p < 0.01).

However, as can be seen in Figure 7, the interactive effects of working memory and comprehension processes tapped by a test item were small in comparison to the main effects of the comprehension processes. That is to say, the findings suggested that text memory and text inference items were easier for students with high working memory capacity when compared with students with low working memory capacity. In other words, students with high working memory capacity had a higher chance of providing correct responses to reading comprehension test items regardless of their type. Interestingly, text memory items were always harder when compared to text inference items regardless of working memory capacity.

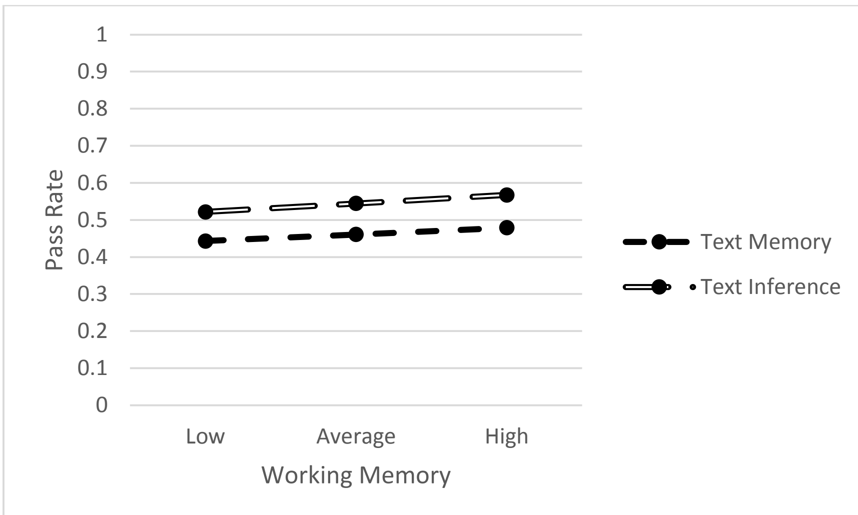


Figure 7. A line plot demonstrating pass rates for text memory and text inference items for low, average and high working memory capacity students.

Lastly, the interaction of referential cohesion with background knowledge indicated that low referential cohesion passages were more difficult for students with low background knowledge but not for students with high background knowledge. Students with high background knowledge seemed to benefit more from more challenging passages (low referential cohesion) as their pass rates on these passages was higher when compared with their pass rates on high referential cohesion passages. That is to say, low referential cohesion was more helpful for students with high levels of background knowledge whereas high referential cohesion was more useful for students with low levels of background knowledge. Interestingly, despite the significant interaction between vocabulary and referential cohesion, pass rates for low and high referential cohesion passages were similar within any given level of vocabulary ability. In other words, low ability students had similar pass rates on low and high referential cohesion passages. The same was true for average and high ability students. This finding suggests that level of referential cohesion did not improve pass rates for students with a given level of ability.

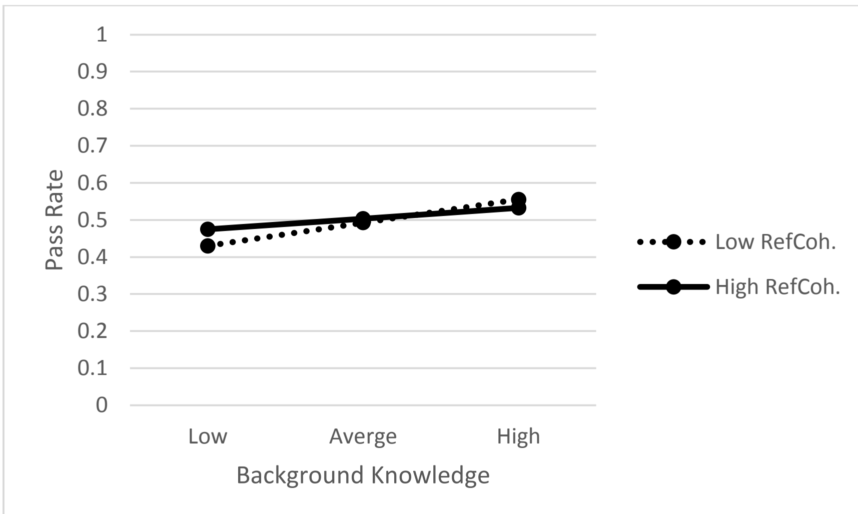


Figure 8. A line plot demonstrating a pass rate for low and high referential cohesion item among students with low, average and high background knowledge. Low RefCoh. = low referential cohesion; High RefCoh. = high referential cohesion.

Discussion

The overarching goal of the current project was to investigate the role of reader characteristics, passage features, comprehension processes tapped by test items, and their interactive effects in explaining variation in reading comprehension test scores. It was expected that reader characteristics would be the most influential in explaining test scores, as stronger cognitive abilities would facilitate understanding of the text regardless of the type and features of the text, or the specific comprehension process measured by the item. At the same time, it was hypothesized that passage features, comprehension processes and the interactive effects of reader characteristics and text features would be less essential in explaining variation in reading comprehension. Consequently, it was expected that variation in responses to reading comprehension test items would be predominantly a consequence of variability in reader abilities and not item or text features as good readers read well across a variety of texts and understand better at both surface and deeper levels of comprehension.

The effects of reader characteristics on reading comprehension. The results revealed that vocabulary and background knowledge were the most influential in explaining variation in reading comprehension test scores regardless of developmental grade span. Readers with high levels of vocabulary and background knowledge had a higher probability of providing correct responses to reading comprehension items when compared to readers with lower levels of vocabulary and background knowledge. The obtained findings were consistent with expectations in so far as constructing a coherent mental representation of the text depends on reader's depth and breadth of knowledge including vocabulary and background knowledge. Better knowledge of word forms and meanings not only influences accuracy and fluency of word reading, which frees up capacity for comprehension, but also helps in deriving meanings of new words and integrating words within an existing text representation (Perfetti, 2007; Vellutino, 2003). Higher level of background knowledge is beneficial in terms of comprehending information conveyed in the text, establishing connections between elements in the text, filling in informational blanks, and integrating textual information with general knowledge (Kendeou & van den Broek, 2007; Kintsch, 1994). Readers with deficient vocabulary or background knowledge have limited ability to understand a variety of texts because their knowledge of words' forms and meanings as well as the overall size of their vocabulary is insufficient for successful inference making and integrating information presented in the text with background knowledge stored in semantic memory (van den

Broek, 2012). Although the present study is only correlational in nature, the results clearly support experimental research which shows that understanding of texts may be improved by learning new words and gaining exposure to new topics. Greater familiarity with a broad spectrum of words and scientific topics enhances learning from texts of variable difficulty. As such, readers become better equipped to understand harder texts involving scientific gorgon. Finally, as we grow older and gain everyday experience we have a somewhat stronger basis for at understanding written language describing that experience in narrative texts, as well as for relating knowledge across domains.

The obtained results also found that: (a) working memory was statistically significant in predicting responses to reading comprehension test items among 7^{th} through 9^{th} graders but not 10^{th} through 12^{th} graders, and (b) reading fluency was important in predicting reading comprehension for 10^{th} through 12^{th} graders, but not 7^{th} through 9^{th} graders. These findings are somewhat surprising as one would expect that working memory and reading fluency would relate to comprehension for all readers. There is no compelling developmental theory to suggest these differential relationships. However, in untimed tests, or tests that allow readers to go back and review information, the effects of working memory and fluency on comprehension may be diminished. The basis for the differential effects of working memory and fluency across the two grade spans is not easily explained on the basis of reading theory, or reading development. Possible explanations include heterogeneity across the two samples of students (i.e., the grade span samples are not strictly comparable) or statistical artifacts. Somewhat higher correlations of working memory with the other reader characteristics were found in grades 10 through 12, and somewhat higher correlations were found between reading fluency and the remaining reader characteristics in grades 7 through 9. These slight, but consistent, differences in the correlation patterns across grades may have added enough statistical noise by increasing standard errors to lead to the difference in findings across grades. Importantly, the coefficients for fluency and working memory are not very different across the two grade spans, and would not be statistically significant if tested as evidenced by the ratio of the difference to the average of the standard errors. Thus, it seems unwise to make too much of this difference in patterns. Rather, the correct conclusion would seem to be that working memory and fluency exerted small effects on test performance in both grade spans.

This difference in findings across grade spans may also be related to heterogeneity in the GMRT test forms. Specifically, text memory and text inference items were equally represented in the GMRT form for grades 7 through 9, but not in the GMRT form for grades 10 through 12 where 80% of test items were text memory items. The inconsistent findings may reflect these differences in the composition of the two test forms. It is possible that the imbalance in the number of text memory and text inference items lead to differential effects of fluency and working memory in the two developmental grade spans. However, arguing against this conclusion is the fact that working memory only weakly interacted with item type, suggesting that working memory was not substantially more important in predicting performance on text memory items as compared to performance on text inferencing items. Thus, while it is possible that form differences contributed to the different roles that predictors played across the grade spans, other findings suggest that the findings across grade spans are more similar than different. Clearly, the two test forms differ both in terms of passage difficulty and in the composition of the questions related to these passages. The precise impact of these form differences on the roles of person and text characteristics in explaining test performance across the two grade spans is unclear and warrants further investigation and replication.

The effects of passage features on reading comprehension. The results revealed that certain passage features explained some of the variation in item difficulties. More specifically, expository passages were more difficult than narrative passages, with pass rates for expository passages below chance levels. The obtained finding was not surprising as there are several factors that also affect the difficulty of expository texts. More specifically, the goal of expository texts is to educate readers about potentially unfamiliar topics which are not related to the reader's everyday experiences (Graesser & McNamara, 2011). As such, these texts are harder to understand because they use less frequent words and may involve informational blanks which cannot be easily filled in with relevant background knowledge (McNamara et al., 1996). That is to say, these texts place higher demands on reader's knowledge as domain-specific knowledge may be required to understand expository texts. Finally, expository texts use less connectives which typically help in clarifying relations between elements in the text (McNamara, Graesser, & Louwerse, 2012). In other words, associations between sentences and ideas have to be inferred as they are often not explicitly stated. Consequently, the use of unfamiliar words and the limited use of connectives may be challenging for some readers and negatively affect understanding of the text.

Variation in the item difficulties of passages for 10th through 12th grades but not 7th through 9th grades was also explained by deep cohesion. More specifically, high deep cohesion passages were less difficult when compared with low deep cohesion passages. This finding was expected as deep cohesion plays an important role in helping a reader associate concepts presented in the text (Graesser & McNamara, 2011). More specifically, high deep cohesion texts are particularly beneficial when the reader tries to understand more challenging material. As such, comprehension of these types of texts is easier because the greater use of connecting words helps to clarify relations between elements in the text. In other words, these texts facilitate comprehension as relations between events and ideas are explicitly stated in the text. Differences in the effects of deep cohesion on reading comprehension test scores in two developmental grade spans might be attributed to greater variability in deep cohesion in passages for grades 10 through 12 than grades 7 through 9. Alternatively, deep cohesion may have influenced the reading comprehension of 10th through 12th graders more than younger students because of the older students increased familiarity with various connectives and their role in conveying meaning.

The results also revealed that traditional measures of text difficulty (i.e., word frequency and sentence length) and referential cohesion were not statistically significant in predicting reading comprehension test scores for both developmental grade spans. These results were a little surprising as previous research has shown that texts with less frequent words, longer sentences and low referential cohesion are more difficult to understand (Graesser, McNamara, Louwerse, & Cai, 2004; Graesser & McNamara, 2011; McNamara, Louwerse, McCarthy, & Graesser, 2010; Perfetti, 1985). One possible explanation is the wide variability in reader differences relative to the variability in text features in this study. It must be kept in mind that the present study employed explanatory item response models to examine performance on a specific measure of reading comprehension designed to differentiate among students within a narrow range of grades. To that end, text features were not manipulated in this study. Rather, natural variation in text features within a chosen test were used to characterize text difficulty. In this regard it is instructive to look at the variability in text features across the two grade spans rather than within grade spans. Such an examination reveals that within each grade span the variability in text features is constrained relative to the variability in text features across both grade spans. Also, the current study found significant effects of text genre, which alone explained a significant proportion of variance in item difficulties. Importantly, expository and narrative texts differ substantially in average word frequency, sentence length and cohesion. As such, the effects of genre may have masked to some degree the effects of word frequency, sentence length and referential cohesion.

The effects of text-reader interactions on reading comprehension. The results for grades 7 through 9 indicated that interactions between referential cohesion and comprehension processes tapped by a test item (i.e., text memory or text inference), referential cohesion and vocabulary, word frequency and vocabulary as well as working memory and deep cohesion were the most important in explaining variation in reading comprehension test scores. High referential cohesion passages increased pass rates on text memory and text inference items. The obtained finding supports the assumption that the overlap of explicit words and concepts in high referential cohesion texts helps in recalling information from the text. More specifically, the overlap stimulates text comprehension through reintroducing the reader to ideas presented earlier in the text (van den Broek et al., 2005; van den Broek & Kendeou, 2008). As such, reactivated concepts become better remembered as they cyclically reappear during reading. Similarly, inferring relations from high referential cohesion texts is easier because these texts involve fewer conceptual gaps. That is to say, inference making is less demanding due to the added support of overlapping information. It is also important to keep in mind that there are many features of texts that make them more or less difficult that were not considered in the present study. For example, informational texts can be made less difficult by incorporation of figures and diagrams. Narrative texts can be made more comprehensible by incorporating summary information. Additionally, the Coh-Metrix framework includes many more dimensions of text difficulty than could be examined in the present study. That is to say, the present study did not attempt to exhaust the universe of possible reader text interactions. Although the present models explained about 70% of the variance in person abilities and 50% of the variance in item difficulties, a substantial portion of both person and item variance remains to be explained.

The obtained results also revealed that readers with high vocabulary had higher pass rates on passages with low or high word frequency and referential cohesion when compared with readers with limited vocabulary knowledge. This finding was expected as more knowledgeable readers are familiar with a wider range of words. Consequently, not only do these readers have fewer problems with comprehending more texts with less familiar words, but they also have a better chance of comprehending harder texts even if these texts do not include cohesion cues to enhance understanding. At the same time, less knowledgeable readers struggle with comprehending texts involving many unfamiliar words and benefit more from explicit word overlap as reduces the amount of informational blanks which need to be inferred from the text (Kendeou & van den Broek, 2007; McNamara et al., 1996). Furthermore, the results suggested that readers with high working memory capacity were more likely to correctly respond to test items for low or high deep cohesion passages. Again, this finding was in concordance with the literature suggesting that students with high working memory capacity possess more resources to store and manipulate currently processed information, allowing them to more accurately make inferences and integrate text information with background knowledge in order to create a coherent representation of the text (Cantor & Engle, 1993; Daneman & Hannon, 2001; Hannon, 2012; Linderholm & van den Broek, 2002; van den Broek, 2012). Together, the results suggest that good readers read well across texts of variable difficulty. Good readers possess more compensatory resources (i.e., semantic knowledge) that they can bring to with the fore in understanding more difficult texts and consequently are better equipped to read a variety of texts regardless of their difficulty. Given the significant roles played by vocabulary and background knowledge in predicting item performance, it stands to reason that educators and practitioners should focus on familiarizing readers with novel words and concepts in order to improve learning experience and maximize students' potential. The pre-teaching of novel concepts and vocabulary is an often recommended strategy for students who are English language learners that has been also argued to confer benefits to all readers, and these findings are consistent with those recommendations.

The results for grades 10 through 12 revealed that students with low background knowledge had higher pass rates on high referential cohesion passages whereas students with high background knowledge had higher pass rates on low referential cohesion passages. Obtained results support an assumption that cohesive cues are particularly helpful for readers with limited knowledge. McNamara and colleagues (1996) have shown that low cohesion texts are more beneficial for readers with high background knowledge whereas high cohesion texts are more beneficial for readers with low background knowledge. More specifically, poorly written texts with many informational blanks require readers to infer information in order to fill in these blanks. As such, readers have to bridge information in order to construct a coherent representation of the text. Because less knowledgeable readers have insufficient background knowledge, filling in information blanks is cumbersome and oftentimes unsuccessful (Hannon & Daneman, 1998; Kintsch, 1998; Ozuru, Dempsey, & McNamara, 2009). Less knowledgeable readers are unable to connect information because they lack information necessary to make inferences. Consequently, they benefit more from texts with overlapping information. On the other hand, more knowledgeable readers are less hindered by texts with low referential cohesion, and may even benefit from reading such texts. For these high ability readers, low referential cohesion stimulates active processing of the text and enhances learning from the text. Active processing not only encourages more knowledgeable readers to form connections between information presented in the text but also reduces redundancy and stimulates learning of the material presented in the text. As such, low referential cohesion helps in deeper understanding of the text and construction of the situation model of the text. In sum, these findings suggest that selection of texts to read will be more comprehensible when selected to reflect students' knowledge. Although it was not examined in the present study, a related concept to knowledge is that of student interest. We were not able to examine students' interest in the topics on which they were being asked to read and answer questions. Research on motivation suggests that student interest will exert a separate influence on comprehension. It is easy to see how interest could be incorporated into the explanatory IRT framework, but that remains a topic for future research.

Finally, across two developmental grade spans the results revealed that text memory items were more difficult when compared with text inference items. This finding was very novel and counterintuitive as one would expect text memory (TM) items to be easier than text inference (TF) items. One possible explanation stems from the relations between features of GMRT passages and the kinds of test items associated with those passages. For instance, in grades 10 through 12 passages that lead to TM items were less story-like as indexed by the narrativity measure (TM = 33, TI = 47) from *Coh-Metrix*, used less frequent words (TM = 3.23, TI = 3.56), had longer average sentence length (TM = 22.95, TI = 19.66), and had lower deep cohesion (TM = 26.75, TI = 61.29). The Lexile measure for passages comprising predominantly text memory items was equal to 1340 whereas the Lexile measure for passages with predominantly text inference items was equal to 1099. Consequently, passages with predominantly text memory items were more difficult. Lower pass rates on text memory items may be attributed to features of passages associated with these items. As such, it was not necessarily true that text memory items were more difficult than text inference items, but that, on average, passages that generated text memory items were more difficult than passages that generated text inference items. Further investigation of this phenomenon is warranted using passages and items that do not confound passage difficulty with item type (i.e., text memory or text inference) in order to better disentangle relations between comprehension processes and reader characteristics.

Implications

There are several important implications which emerged from the present study. First of all, understanding and learning from the text may be improved by learning new words and familiarizing oneself with novel concepts. As we learn new words, we become more experienced and gain better understanding of what authors are trying to tell as. As such, improving vocabulary affords readers the opportunity to extend their reading to ever more complex text. It is clear that word frequency affects text difficulty, and knowledge of word meanings mitigates that effect on comprehension. Second, good readers read well across a variety of texts as they are more equipped to deal with challenges introduced by the texts. However, it is important to also keep in mind that readers' learning experiences are maximized when they are forced to actively think about ideas presented in the text. In other words, texts should not be too easy as simple texts do not encourage the reader to think about conveyed information. Third, reading and learning as measured on an achievement test may not be the same as reading and learning from books and/or more extended reading, such as magazine and news articles. Specifically, even though it is well established that children learn through reading, it is not obvious if the processes involved in reading texts used in operational tests and other standard reading materials are comparable, and whether they are similarly influenced by reader and text characteristics. The National Assessment of Educational Progress (NAEP) has a very complex framework for the design of reading passages that closely parallels the kinds of reading that students are asked to perform in school and outside of school.

The current project, looked at understanding and learning from a more typical standardized achievement test. We found that understanding and learning of the text was primarily affected by readers' semantic knowledge and was not especially dependent on how the texts were written. It would be interesting to see if these findings can be replicated when readers read longer texts, or are forced to engage in more complex processing of texts, such as in the NAEP. It is possible that longer texts may require a different level of engagement during reading. Finally, it would be interesting to see whether these findings will hold on the new assessments being designed to coincide with the Common Core State Standards as these assessments intend to determine what students know and are capable of doing in relation to more complex standards for college and career readiness. As such, it will be important to replicate the current study with a broader variety of tests and test items than is afforded by a single comprehension assessment, and the new Common Core assessments and the released items from the NAEP would seem important assessments to consider in such a replication.

Limitations

The limitations of the current study include inadequate length and insufficient number of passages, lack of common items between GMRT passages for grades 7 through 9 and 10 through 12, as well as ignoring the nesting of items within passages. All passages were rather short as they included less than 200 words. As such, cohesion measures assessed with the Coh-Metrix Text Easability Assessor may have been less meaningful. Furthermore, variability of text features may have been limited because of the insufficient number of passages per developmental grade span. Because the passages were not experimentally controlled, but naturally occurring in the context of the GMRT, not only was it impossible to manipulate passage features, but it was also impossible to control the relationships among features of passages. For instance, low referential cohesion passages were found to be less difficult than high referential cohesion passages as measured by the average word frequency. Specifically, the low referential cohesion passages had an average word frequency of 3.42 compared to 3.35 for the high referential cohesion passages. Similarly, the average sentence length was equal to 16.77 and 24.15 for low and high referential cohesion passages, respectively. Thus, the effects of cohesion on text difficulty were mitigated by the concomitant effects of word frequency and sentence length, which were found to be operating to make the low referential cohesion passages easier. These results coincide with findings from Graesser and McNamara (2011) who also found that high referential cohesion texts may have side effects at the surface code level including increased word frequency, sentence length and syntactic complexity. As such, conclusions regarding effects of passage features on reading comprehension may not generalize beyond the naturally occurring passages on the GMRT, and cannot be taken to equal the causal effects of these passage features on comprehension.

Furthermore, it was not possible to jointly model the two forms of the test and place all parameters across grade spans on a common scale because there were no common items across the two-forms of the GMRT, and the study design did not otherwise provide a means for equating between the grade spans. Modeling each form separately limited comparison of findings between the two developmental grade spans. As such, discussion of differences in the effects of reader and text characteristics on reading comprehension between the two developmental grade spans should be treated with caution. It was difficult to determine whether the obtained patterns of results for the two developmental grade spans were a function of developmental changes or artifacts related to the grade-specific GMRT passages and the corresponding items. Disentangling patterns of results for the two developmental grade spans may be possible if the Lexile measure is used to equate comprehension scores across the two grade blocks as this measure is interval scaled and provides some basis for equating comprehension scores across grades. However, accomplishing this equating in the explanatory IRT framework is a separate research project in its own rite.

Another limitation of the statistical modeling in the current study is the failure of the models to address clustering of items within passages. Ignoring the nesting of items within passages may have affected estimation of passage effects and item effects nested within passages, and certainly the standard errors for both person and item effects. More specifically, estimation of item effects and their standard errors may have been biased as passages were variable in terms of text features. That is to say, item responses across different items for the same passage are not independent, but the model treats them as independent. As such, items do not carry completely independent information about the effects of text or person characteristics on reading comprehension. Although ignoring clustering tends to underestimate standard errors, the potential biasing effects of clustering on regression parameter estimates are more difficult to predict and may result in either over- or under-estimation of the size of regression parameters.

Despite these limitations, the present study is one of only a few studies investigating the effects of text reader interactions on reading comprehension utilizing explanatory item response models. Examination of these effects using explanatory item response models carries important implications as it provides meaningful insights related to the effects of dynamic relations between reader and text characteristics on understanding of the text. In other words, the application of explanatory item response models in the context of the current study allowed the modeling of individual differences in reading comprehension test scores as a function of reader characteristics, text features and their interactions. Importantly, this approach revealed possible sources of text difficulty in GMRT that differentially affect the comprehension of readers depending on characteristics of the reader. As such, the research showed how students with different levels of various abilities perform on the operational test with correlated text features and suggested ways in which to intervene for children based on their profiles of skills and abilities.

Conclusions

The findings of the present study are particularly informative as they reveal the interplay of reader characteristics, text features and comprehension processes. They show how relations between reader and text affect within-person variability in performance on test items. More specifically, they explain why certain types of texts may be more challenging for readers with particular skills and abilities. As such, they demonstrate why less skilled readers are more challenged by texts with specific features. Consequently, they help in identifying reader strengths and weaknesses, and provide a basis for predicting how students will perform on particular types of items addressing particular kinds of texts on future examinations of reading comprehension. As explanatory models, they also provide a basis for creating interventions that can be used to change that performance. Continued exploration of text-reader interactions will further improve our understanding of the mechanisms involved in reading comprehension and how best to explain and improve the reading comprehension of individual students.

References

Anderson, R.C. (1972). How to construct achievement tests to assess comprehension. *American Education al Research Association, 42(2),* 145-170.

Baddeley, A.D. (1998). Recent developments in working memory. *Current Opinion in Neurobiology, 8*, 234-238.

Barnes, M.A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inference in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*, 216-241.

Bates, D., Maechler, M., & Dai, B. (2008). The lme4 Package version 0.999375-26. http://cran.r-project.org/web/packages/lme4/lme4.pdf/.

de Boeck, P. & M. Wilson. (2004). *Explanatory item response models: A Generalized linear and nonlinear approach*. New York, NY: Springer.

Cantor, J., & Engle, R.W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1101-1114.

Carver, R.P. (1990). *Reading Rate: A review of research and theory.* New York, NY: Academic Press.

Conklin, H.M., Luciana, M., Hooper, C.J., & Yarger, R.S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology, 31*, 103-128.

Christopher, M.E., Miyake, A., Keenan, J.M., Pennington, B., DeFries, J.C., Wadsworth, S.J., …Olson, R.K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: a latent variable analysis. *Journal of Experimental Psychology: General, 141(3),* 470-88. doi: 10.1037/a0027375.

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.

Davis, E., Barnes, M., and Denton, C. (2013). Gates-MacGinitie Reading Tests – Background Knowledge (GMBK). Unpublished PACT assessment task descriptions. TIMES, Houston.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102(2),* 211-245.

Fischer, G.H., & Molenaar, I.W. (1995). *Rasch models: Foundations, recent developments, and applications*. New York, NY: Springer.

Fletcher, J.M., Francis, D.J., O'Malley, K., Copeland, K., Mehta, P., Caldwell, C., …Vaughn, S.R. (2009). Effects of a bundled accommodations package on high stakes testing for middle school students with reading disabilities. *Exceptional Children*, 75, 412-428.

Fletcher, J.M., Reid Lyon, G., Fuchs, L.S., & Barnes, M. (2007). *Learning Disabilities: From Identification to Intervention.* New York, NY: Guilford Press.

Gathercole, S.E., & Alloway, T.P. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40,* 177-190.

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

Graesser, A.C., McNamara, D.S., & Kulikowich, J.M. (2011). Coh-Metrix: Providing multilevel analysis of text characteristics. *Educational Researcher, 40,* 223-234.

Graesser, A.C., D.S. McNamara, & Louwerse, M.M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text?. In A. Polselli & C.E. Sweet. Snow *Rethinking Reading Comprehension*, (pp. 82-99) New York, NY: Guilford Press.

Graesser, A.C., McNamara, D.S., Louwerse, M.M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36(2)*, 193-202.

Graesser, A.C., & McNamara, D.S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science, 3(2),* 371-398.

Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly, 47(2),* 125-152.

Hannon, B. & Daneman, M. (1998). Facilitating knowledge-based inferences in less-skilled readers. *Contemporary Educational Psychology, 23,* 149-172.

Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology, 93(1),* 103-128.

Hannon, B., & Daneman, M. (2009). Age-related changes in reading comprehension: an individual-differences perspective. *Experimental Aging Research, 35(4),* 432-456.

Hitch, G.J., Towse, J.N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130,* 184-198.

Kemps, E., De Rammelaere, S., & Desmet, T. (2000). The development of working memory: Exploring the complementarity of two models. *Journal of Experimental Child Psychology, 77,* 89-109.

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

Kintsch, W. (1979). On modeling comprehnesion. *Educational Psychologist, 14,* 3-14.

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

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49(4),* 294-303.

Kintsch, W. (1998). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95(2),* 1101-1117.

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes, 39(3),* 125-128.

Kintsch, W. (2012). Psychological models of reading comprehnesion and their implications for assessment. In J.P. Sabatini, E. Albro, & R.T. O'Reilly, *Measuring Up,* (pp. 21-38) Lanham, MA: R&L Education.

Linderholm, T., Everson, M.G., van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2001). Effects of causal text revision on more or less-skilled readers' comprehension of easy and difficult texts. *Cognition and Instruction, 18,* 525-556.

Linderholm, T., Virtue, S. Tzeng, Y., & van den Broek, P. (2004). Fluctuations in the availability of information during reading: capturing cognitive processes using the Landscape Model. *Discourse Processes, 37(2),* 165-186.

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.

Lorch, R.F., & van den Broek, P. (1997). Understanding reading comprehension: Current and future contributions of cognitive science. *Contemporary Educational Psychology, 22,* 213-246.

Lynch, J.S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on- and off-line inferences about characters' goals. *Cognitive Development, 22(3),* 323-340.

MacGinitie, W.H., MacGinitie, R.K., Maria K., & Dreyer, L. (2000). *Gates-MacGinitie Reading Tests,*

4th edition. Chicago: Riverside Publishing.

McNamara, D.S., (2001). Reading both high and low coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55,* 51-62.

McNamara, D.S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22,* 247-288.

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch. W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14(1),* 1-43.

McNamara, D.S., Louwerse, M.M., McCarthy, P.M., & Graesser, A.C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes, 47(4)*, 292-330.

McNamara, D.S., A.C. Graesser, & Louwerse, M.M. (2012) Sources of Text Difficulty: Across the Ages and Genres, In J.P. Sabatini, E. Albro, & R.T. O'Reilly, *Measuring Up,* (pp. 89-119) Lanham, MA: R&L Education.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28*, 292–312.

National Center for Education Statistics (2013). *The Nation's Report Card: A First Look: 2013 Mathematics and Reading* (NCES 2014–451). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Ozuru, Y., et al., *Where's the difficulty in standardized reading tests: The passage or the question?* Behavior Research Methods, 2008. 40(4): p. 1001-1015.

Ozuru, Y., K. Dempsey, & McNamara, D.S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19(3)*, 228-242.

Perfetti, C.A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehnesion. *Scientific Studies of Reading, 11*, 357-383.

Perfetti, C.A., & Adolf, S.M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning, In J.P. Sabatini, E. Albro, & R.T. O'Reilly, *Measuring Up,* (pp. 3-20) Lanham, MA: R&L Education.

Rapp, D., & van den Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science, 14(5),* 276-279.

Raven, J., Raven, J.C., & Court, J.H. (1998). *Coloured Progressive Matrices 1998 edition.* Oxford: Oxford Psychologists Press.

Reise, S.P. Ainsworth, A.T., & Haviland, M.G. (2005). Item Response Theory. Fundamental, applications, and promise in psychological research. *Current Directions in Psychological Science, 14(2),* 95-101.

Schnick, T., & Knicklebine, M. (2006-2007). *Using the Lexile Analyzer foreEducators and media specialists,* Durham: MetaMetrics, Inc.

Snow, C.E. & Polselli Sweet, A. (2003). Reading for comprehnesion. In A. Polselli & C.E. Sweet. Snow *Rethinking Reading Comprehension,* (pp. 1-11) New York, NY: Guilford Press.

Tilstra, J., McMaster, K., van den Broek, P., Kendeou, & Rapp, D. (2009). Simple but complex: components of the simple view of reading across grade levels. *Journal of Research in Reading, 32(4),* 383-401.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency (TOWRE).* Austin, TX: Pro-Ed.

Tzeng, Y., van den Broek, P., Kendeou, P., & Chengyuan, L. (2005). The computational implementation of the Landscape Model: Modeling inferential processes and memory representations of text comprehension. *Behavior Research Methods, 37(2),* 277-286.

van den Broek, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39(3),* 299-316.

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science, 328(453),* 453-456.

van den Broek, P. (2012). Individual and developmental differences in reading comprehension: Assessing cognitive processes and outcomes, In J.P. Sabatini, E. Albro, & R.T. O'Reilly, *Measuring Up*, (pp. 39-58) Lanham, MA: R&L Education.

van den Broek, P., & Espin, C.A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review, 41(3),* 315-325.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: the role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22(3),* 335-351.

van den Broek, P., Lorch, E.P., & Thurlow, R. (1996). Children's and adult's memory for television stories: The role of causal factors,story-grammar categories and hierarchical level. *Child Development, 67,* 3010-3029.

Van den Broek, P., Rapp, D.N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, 39, 299–316.

van den Noortgate, W., de Boeck, P., & Meulders, M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics, 28(4)*, 369-386.

Vellutino, F.R. (2003). Individual differences as sources of variability in reading comprehension in elementary children, In A. Polselli & C.E. Sweet. Snow *Rethinking Reading Comprehension,* (pp. 51-81) New York, NY: Guilford Press.

von Davier, M., & Carstensen, C.H. (2006). *Multivariate and mixture distribution Rasch models: Extensions and applications*. New York, NY: Springer.

Woodcock, R. W., McGrew, K., & Mather, N. (2007). *Woodcock-Johnson Tests of Cognitive Skills*.
3rd edition. Itasca, IL: Riverside Publishing.

Yeari, M., & van den Broek, P. (2011). A cognitive account of discourse understanding and discourse interpretation: The Landscape Model of reading. *Discourse Studies, 13(5)*, 635-643.

Appendix A

I. MATHEMATICAL FORMULATION OF THE RASCH MODEL

1. The basic formulation of the Rasch model

$\eta_{pi} = \beta_i + \theta_p,$ $\theta_p \sim N(0, \sigma_\theta^2)$

with η_{pi} as a linear component for each pair of a person p and an item i;

β_i as the person parameter;

θ_p as the item difficulty, where θ_p is normally distributed with the mean of 0.

2. The exponential form of the Rasch model

$$\eta_{pi} = \exp(\theta_p - \beta_i)$$

3. The Rasch model expressed as the ratio of the success probability to the failure probability.

$\eta_{pi} = \log(\frac{\pi_{pi}}{1 - \pi_{pi}})$

with π_{pi} as the probability of success and $1 - \pi_{pi}$ as the probability of failure.

4. The Rasch model expressed as the ratio of the success probability to the failure probability as the ratio of a person's ability to the difficulty of the item.

$\frac{\pi_{pi}}{(1 - \pi_{pi})} = \frac{\exp(\theta_p)}{\exp(\beta_i)}$

II. A SIMPLIFIED MATHEMATICAL FORMULATION OF THE DISSERTATION MODEL

(excluding interactions)

$$\eta_{pi} = \underbrace{\boxed{\sum_{j=1}^{J} \theta_j Z_{pj} + \ \varepsilon_p}}_{\text{PERSON CONTRIBUTION}} - \underbrace{\boxed{\sum_{k=1}^{K} \beta_k X_{ik} + \ \varepsilon_i}}_{\text{ITEM CONTRIBUTION}}$$

$$\varepsilon_p \sim N(0, \sigma^2{}_{\varepsilon})$$
$$\varepsilon_k \sim N(0, \sigma^2{}_{\varepsilon})$$

with η_{pi} as a linear component for each pair of a person *p* and an item *i*, and θ_p and β_i (from equation 1) formulated as linear regression equations.

A. Linear regression equation representing person contribution

j, j = 1, …, J as an index for the person covariates;

θ_j as a fixed regression weight of person property j;

Z_{pj} as a person predictor, value of person *p* on person property *j*;

ε_p as a residual person variance, where ε_p is normally distributed with the mean of 0.

B. Linear regression equation representing item contribution

k, k = 1, …, K as an index for the item covariates;

β_k as a fixed regression weight of item property *k*;

X_{ik} as an item predictor, value of item *i* on item property *k*;

ε_i as a residual item variance, where ε_k is normally distributed with the mean of 0.

Appendix B

Table 8. Bivariate correlations between reader characteristics by two grade spans (7 through 9 and 10 through 12)

	1	2	3	4	5
GRADES 7 - 9					
1. Word reading	-				
2. Reading fluency	0.50***	-			
3. Vocabulary	0.51***	0.26***	-		
4. Background knowledge	0.43***	0.24***	0.70***	-	
5. Working memory	0.32***	0.25***	0.32***	0.28***	-
GRADES 10 - 12					
1. Word reading	-				
2. Reading fluency	0.49***	-			
3. Vocabulary	0.54***	0.31***	-		
4. Background knowledge	0.51***	0.33***	0.76***	-	
5. Working memory	0.28***	0.32***	0.27***	0.22***	-

Note: *p < .05 **p < .01 ***p < .001