Published in final edited form as:

J Res Educ Eff. 2016 April 2; 9(2): 153–172. doi:10.1080/19345747.2015.1111482.

Cognitive Difficulties in Struggling Comprehenders and their Relation to Reading Comprehension: A Comparison of Group Selection and Regression-Based Models

Marcia A. Barnes¹, Karla Stuebing², Jack M. Fletcher², Amy Barth³, and David Francis²

¹Department of Special Education & Meadow's Center for Preventing Educational Risk, University of Texas, Austin, Texas

²Department of Psychology, University of Houston, Houston, Texas

³Department of Special Education, University of Missouri, Columbia, Missouri

Abstract

Difficulties suppressing previously encountered, but currently irrelevant information from working memory characterize less skilled comprehenders in studies in which they are matched to skilled comprehenders on word decoding and nonverbal IQ. These “extreme” group designs are associated with several methodological issues. When sample size permits, regression approaches permit a more accurate estimation of effects. Using data for students in grades 6 to 12 ($n = 766$), regression techniques assessed the significance and size of the relation of suppression to reading comprehension across the distribution of comprehension skill. After accounting for decoding efficiency and nonverbal IQ, suppression, measured by performance on a verbal proactive interference task, accounted for a small amount of significant unique variance in comprehension (less than 1%). A comparison of suppression in less skilled comprehenders matched to more skilled comprehenders (48 per group) on age, word reading efficiency and nonverbal IQ did not show significant group differences in suppression. The implications of the findings for theories of reading comprehension and for informing comprehension assessment and intervention are discussed.

Cognitive models of reading comprehension propose that cognitive resources, such as working memory and associated processes such as suppression of irrelevant information from working memory are important for supporting the construction of an accurate representation of the situation that the text describes. Furthermore, variability in these and other cognitive abilities have been proposed as important sources of individual differences in reading comprehension (Gernsbacher, 1996; Gernsbacher & Faust, 1991; Just & Carpenter, 1992; van den Broek, Rapp, & Kendeou, 2005). What seems of critical importance for comprehension theory, assessment and instruction is to understand which of the many factors that are related to reading comprehension are actually *integral* to reading comprehension (Perfetti & Adlof, 2012). As such, determining the relative importance of comprehension-related factors to comprehension involves not only establishing whether

there is a significant relation between these factors and comprehension, but also whether the *size* of that relation is large enough to be of practical significance for assessment and intervention. The answer to these questions requires the use of methods that can adequately address questions of significance and magnitude of effect. This paper uses regression based models to address these questions regarding the relation of suppression and reading comprehension.

Suppression and Reading Comprehension

Models of higher cognition, including reading comprehension, posit an important role for capacity-limited executive abilities, particularly working memory and inhibition (Geary, 2013). Inhibitory processes are central to all models of executive function (Baddeley, 2000; Engle & Kane, 2004; Miyake & Shah, 1999). In Engle's model, controlled attention is considered to be a domain general capacity that determines the contents of working memory by actively maintaining context-relevant information and suppressing context-irrelevant representations in working memory (Engle & Kane, 2004). In the seminal latent variable study of Miyake and his colleagues (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), evidence was found for three moderately related, but also somewhat separable aspects of executive functioning - mental shifting, updating, and inhibition.

Although inhibition and suppression are sometimes used interchangeably, suppression is best considered as part of a family of inhibitory processes (Friedman & Miyake, 2004). A latent variable analysis of inhibition-related functions, for example, showed overlap between the constructs of response inhibition (e.g., measured by tasks such as the Stop Signal Response Time or SSRT) and resistance to distractor interference (e.g., ability to ignore conflicting information as required in Flanker paradigms). A separate inhibition-related construct, resistance to proactive interference, emerged from these analyses. Proactive interference is tapped by memory paradigms in which information that is no longer relevant is not able to be suppressed and is mistakenly recalled; in models of reading comprehension, discussed below, this type of inhibitory function has been referred to as *suppression*.

Some cognitive process models of comprehension include the operation of domain general cognitive abilities such as working memory and suppression in describing how the representation of the situation described by text is constructed. The Structure Building Framework (Gernsbacher, 1996) posits a central role for suppression in successful comprehension. According to this model, readers construct mental structures, the "building blocks" of which are memory nodes. It is the modulation of these nodes that result in the dynamic building of coherent mental structures or representations that characterize comprehension. In this comprehension model, the building of these mental structures requires two complementary mechanisms, *enhancement*, which boosts the activation of memory nodes that have been previously laid down when the incoming information in the text is relevant to the structure being built, and *suppression*, which dampens the activation of memory nodes associated with information that is no longer as necessary to the structure being built. When information is perceived as not being sufficiently relevant to the ongoing building of a mental structure, the reader shifts to develop a new substructure. Text information in mental substructures that have been shifted away from becomes less

accessible. In the Structure Building Framework, it is proposed that less skilled comprehenders shift to develop new substructures too often. The cognitive mechanism behind this “over shifting” is thought to be inefficient suppression.

Similar to domain general views of working memory (e.g., Engle & Kane, 2004), Gernsbacher and her colleagues refer to these mechanisms for comprehension as also being domain general because they operate similarly whether when one is understanding oral or written language or pictorial sequences of events and in a variety of other verbal and nonverbal situations (Gernsbacher & Faust, 1991; Gernsbacher, Varner, & Faust, 1990).

Individual Differences in Suppression

Two types of paradigms are used in the reading comprehension literature to study the relation of suppression and comprehension. Studies using ambiguity resolution paradigms measure how quickly readers can suppress contextually irrelevant word meanings, for example, the “money” meaning of “bank” when reading “He sat on the river bank”. These studies suggest that adults and children who are poor comprehenders have less efficient suppression; that is, the context-irrelevant meaning of the ambiguous word (i.e., the money meaning of bank in the example above) is actively maintained in memory for longer than it should be in poor comprehenders compared to good comprehenders (e.g., Barnes et al., 2004; Gernsbacher & Faust, 1991). Difficulties in suppression rather than enhancement are presumed to differentiate good from poor comprehenders because both groups activate both meanings of “bank” immediately after reading the word regardless of the context, but 1000 msec. later, the contextually-irrelevant meaning of bank remains activated only for poor comprehenders; it is presumed to have been suppressed by good comprehenders. Thus, previously activated information that is no longer relevant remains active and interferes with the ongoing construction of the text representation for poor comprehenders. Because this suppression effect occurs about a second after a word is read (but not within about the first 500 msec.), suppression is considered to be a controlled or active process.

Directed forgetting tasks are also used to measure suppression (Wilson & Kipp, 1998). In these tasks, also called proactive interference tasks, the mistaken recall of to-be-forgotten information (i.e., intrusion errors in recall) is taken as an indicator of difficulties in suppression. Findings from studies measuring intrusion errors suggest that adults and children who are poor comprehenders are more likely than good comprehenders to recall words that are no longer relevant or which participants have been directed to forget (Caretto, Cornoldi, De Beni, & Romano, 2005; De Beni & Palladino, 2000; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Pimperton & Nation, 2010). In one study (Borella, Caretti, & Pelegrina, 2010), good and poor comprehenders differed in performance on proactive interference and working memory tasks, but not on response inhibition and distractor interference tasks, lending further support for the idea that inhibitory functions related to suppression are separable from those related to response inhibition and distractor interference (Friedman & Miyake, 2004). Similar to the ambiguity resolution studies, suppression in the proactive interference paradigm is also purported to reflect an active, controlled process.

There has been some disagreement about whether suppression difficulties experienced by poor comprehenders are specific to language-based processing or whether they are domain general; that is, whether the difficulty resides in a basic inhibitory mechanism that affects processing across domains and materials as suggested by Gernsbacher (1996), or whether the difficulty is only in the verbal domain. Recently, Pimperton and Nation (2010) found evidence for domain specificity of suppression deficits such that poor comprehenders differed from good comprehenders only on verbal versions of their proactive interference tasks.

In both the ambiguity resolution studies and the proactive interference studies reported above, good and poor comprehender groups are created by matching, sometimes on a case-by-case basis, for age, word reading accuracy or reading fluency, and sometimes for nonverbal cognitive abilities. The logic behind this matching design is that groups that differ significantly in comprehension, but who are matched for word reading and nonverbal IQ, can then be compared on another variable hypothesized to be important for comprehension, such as suppression. If a group difference emerges on the target variable, an inference is made that this variable is a good candidate for producing individual differences in comprehension given that other person characteristics such as word reading ability and nonverbal abilities have been “equated” through the matching process.

Less common are studies that test these relations across a range of reading skill using regression-based designs. Two recent studies of this type have reported somewhat different findings for different measures of inhibitory functions and reading comprehension. Arrington, Kulesz, Francis, Fletcher, & Barnes (2014), in a study of middle school and high school students, measured response inhibition using the Stop signal reaction time task (SSRT) and suppression using the Pimperton and Nation (2010) verbal proactive interference task. In path analyses, they found that performance on the verbal proactive interference task was directly related to reading comprehension, while performance on the SSRT was directly related to reading decoding. In a study of both younger students and adolescents by Christopher et al. (2012), a latent inhibition construct comprised of response inhibition measures (SSRT and a continuous performance task) did not account for unique variance in reading comprehension; however, neither of the measures used in that study assessed inhibitory functions in the verbal domain, nor was proactive interference measured.

In sum, findings from studies using small groups of good and poor comprehenders, and one large study using path analysis, suggest that difficulties in suppression tapped by verbal proactive interference tasks is associated with comprehension (Arrington et al., 2014; Borella et al., 2010; Pimperton & Nation, 2010). However, these studies did not assess the relation of suppression and reading comprehension controlling for general task demands in proactive interference paradigms (i.e., cued recall abilities when suppression is not required). For example, in the Pimperton and Nation study (2010) separate analyses were conducted for interference trials (trials on which participants were directed to forget something they had heard before) and no-interference trials (trials on which participants were not directed to forget what they had previously heard). Arrington et al. (2014) did not include performance on the no-interference trials in their path analysis as a means of controlling for general cued recall abilities. Thus, neither extreme group nor regression

studies have addressed the unique influence of suppression on reading comprehension. Furthermore, there are several methodological issues that arise when using extreme group designs as discussed below.

Methodological Issues in Determining the Relevance of Cognitive Processes for Reading Comprehension

Group-based studies with relatively small numbers of participants with extreme scores on the dependent variable of interest are often used in individual difference studies of comprehension. Although these studies are appropriate in early stages of research for generating hypotheses about the sources of individual differences in comprehension, it has been argued that this design is lacking in internal, external, construct and statistical conclusion validity (Jackson & Butterfield, 1989). Beyond the initial stages of investigation, other approaches are likely to lead to more informative estimates of effects (Cain, Oakhill & Bryant, 2000; Preacher, Rucker, MacCallum & Nicewander, 2005). In short, strong causal inferences cannot be made from these studies because they are observational and correlational rather than experimental (Jackson & Butterfield, 1989). Additionally, extreme group designs lead to difficulties in interpreting effect sizes due to effect size inflation (Preacher et al., 2005). The reduction in standard deviation brought about by selecting homogeneous and extreme groups exaggerates the mean difference between groups when an effect size is calculated.

This procedure also lacks strong external validity because of the use of cut-points to determine categories or classes of readers. The assumption is made that classes of individuals exist and that the cut-points validly capture group membership, which is questionable because of the dimensional nature of the reading attributes (Brannum-Martin, Stuebing, Fletcher, & Francis, 2012; Ellis, 1984). External validity is also poor because matching of participants requires the creation of a highly artificial sample through deletion of participants with no match in the sample. The more characteristics to be matched, the more participants must be removed from the sample. Paradoxically, the better the matching, the less representative the sample is of the population that one hopes to make inferences about. This is a particular problem when group creation is done on the basis of extreme scores, which tend to have the largest measurement errors (Campbell & Kenny, 1999).

There is also potential for model misspecification in the extreme groups design. The implicit assumption of this design is that the functional relation of two variables is linear across the continuum of ability. This is not necessarily the case and cannot be explored when only extreme cases are investigated through group comparisons. This question can be explored via regression analysis on cases selected to represent a wide range of ability levels. Note that both extreme group comparisons and regression models are vulnerable to other sorts of model misspecification related to the selection of matching variables in the group design or covariates in the regression design.

Finally, extreme group studies are often described as experiments rather than observational or correlational studies, which lends strength to the idea that they have more validity than the observational studies they represent. The language suggests that participants are assigned to

groups when they are only classified into groups on the basis of measures. There is no intervention or experimental manipulation to cause a difference in the dependent variable. There is only a measure that has been selected as outcome. The statistical tests (F 's and t 's) typically carried out on these group designs assume that the outcome would be unrelated to the grouping variable in the absence of an effective experimental treatment. However, in these observational studies, this is not plausible and the test is not really valid. In observational studies measures are typically correlated with each other to some degree and the only reasonable approach is to estimate the strength of the correlation. To reiterate, selecting extreme groups exaggerates the strength of the correlation by comparing the difference between groups based on a reduced within group variability constrained by the parameters used to select participants.

Current Study

The primary goals of the current study, were to use a regression-based approach to investigate the relation of suppression and reading comprehension, the magnitude of this relation, and whether the relation is invariant across different levels of comprehension. We also sought to compare findings from this regression-based approach to those using an extreme groups design in which we matched participants for age, grade, word reading, and nonverbal IQ, similar to the approach taken in other studies of good and poor comprehenders. With the exception of the age of the participants, the group comparison most closely resembled the Pimperton and Nation (2010) study, but it was not an exact replication. For the regression-based analyses, a large sample was selected to allow estimation of effects with precision and we included participants across the range of ability in reading comprehension, which allowed us to determine the functional relation between suppression and reading comprehension and to guard against the possibility that this functional relation is non-linear. We also tested the specific contribution of suppression to reading comprehension by controlling for cued recall abilities in the absence of the need to suppress information that one has been directed to forget.

Our research questions were as follows: 1) Is there a relation of suppression and reading comprehension in a large sample of adolescent readers using a regression-based design that assesses this relation across the distribution of reading comprehension performance? 2) Is this relation invariant across reading comprehension performance? 3) What is the magnitude of this effect? 4) Are findings from the regression-based analysis and an extreme groups approach similar?

Method

Participants

Sampling Strategy—In the analyses reported below only those participants who had values for every variable were included ($n = 766$). The larger sample from which this group was derived represented 1,765 students in grades 6 through 12 from mainstream classrooms in four school districts within a large southwest city metropolitan area. Selection was based, in part, on performance on the previous year's administration of the Texas Assessment of Knowledge and Skills (TAKS), the state reading accountability test and a reliable and valid

measure of reading comprehension (Cirino et al., 2013). Students were randomly selected from subgroups who met or did not meet benchmark criteria on the TAKS. We randomly selected students within each class, but oversampled students with poor TAKS performance, so that 47% of the sample passed and 53% did not pass. Students were excluded from participation if their school identified them as Limited English Proficient (LEP), if their reading instruction or English Language Arts instruction was provided by a LEP teacher, or if they had a significant disability (e.g., intellectual-cognitive disabilities, severe behavioral disabilities, or autism).

Students who consented were screened on word decoding and general intelligence. Students who scored at or above the 20th percentile on the Woodcock-Johnson III Tests of Achievement, Letter Word Identification subtest (WJIII-LWID; Woodcock, McGrew, & Mather, 2007) were eligible to continue because we wanted to have a sample whose comprehension difficulties were not primarily related to very low word reading. Students also had to have a verbal and/or fluid intelligence score at or above 70, as determined by the Kaufman Brief Intelligence Test – 2 (KBIT-2; Kaufman & Kaufman, 2004) in order to rule out intellectual disability. In total, 166 students refused consent and 411 were disqualified due to low word reading. Students who passed the screening measures ($n = 1,352$) were then tested on a larger battery.

The main reason for excluding students from the analyses in this study is that they did not have KBIT-2 Matrices scores. KBIT-2 Matrices scores were missing by design; that is, students were given either KBIT-2 Matrices or KBIT-2 Verbal Knowledge to provide a measure of general cognitive development to exclude for intellectual disabilities. 787 students had KBIT-2 Matrices scores; however, 21 students with KBIT-2 scores were missing scores on other measures used in this study, resulting in a final sample of 766.

Comparisons of reading scores and demographics of participants with and without KBIT-2 Matrices scores were conducted. Participants with and without KBIT-2 Matrices scores did not differ on any of the word reading or reading comprehension measures (Gates MacGinitie Reading Test, Reading Comprehension subtest [GMRT-RC], $t(1204) = .94, p = .35$; Test of Word Reading Efficiency Phonemic Decoding [TOWRE-PD], $t(1203) = .88, p = .38$; Test of Word Reading Efficiency Sight Words [TOWRE-SW], $t(1204) = 1.59, p = .11$; WJIII-LWID $t(1315) = .13, p = .90$) or on demographic variables (Gender $\chi_{(1)} = 2.5293, p = .12$; Race $\chi_{(5)} = 2.41, p = .79$), but there was a small effect for Free and reduced lunch status (Free and reduced lunch $\chi_{(1)} = 4.06, p = .044$). Those with KBIT-2 scores and full data did not differ from those with KBIT-2 scores who were missing one or more scores on other variables used in this study (GMRT-RC, $t(768) = -1.06, p = .29$; TOWRE-PD, $t(784) = .49, p = .87$; TOWRE-SW, $t(785) = .38, p = .71$; $p = .62$; WJIII-LWID $t(785) = 1.38, p = .17$; Gender $\chi_{(1)} = .28, p = .61$; Race $\chi_{(5)} = 2.79, p = .73$; Free and reduced lunch $\chi_{(2)} = 2.06, p = .36$).

Measures

Texas Assessment of Knowledge and Skills (TAKS; Texas Education Agency)

—The TAKS is a group administered, criterion-referenced assessment of reading comprehension, with grade specific forms. The TAKS requires students to read both expository and narrative passages and answer comprehension questions, and is designed to

assess factors such as critical thinking, use of strategies, and analysis. Internal consistency for the 2010 TAKS for grades 7-12 ranges from .73 to .89, and from .87-.89 for the 2011 TAKS. The TAKS was used as a screening measure as discussed above.

Woodcock-Johnson III Tests of Achievement, Letter Word Identification subtest (WJIII-LWID; Woodcock et al., 2007)—The Letter Word Identification subtest requires children to read real words that vary in difficulty. The test is individually administered, and had a mean reliability coefficient of .91 across our sample. WJIII-LWID was both used as a screening measure as discussed above and in the regression analyses.

Decoding efficiency—The Sight Word Reading Efficiency and Phonemic Decoding subtests of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) were used to assess word reading efficiency in the regressions and to match groups in the group-based comparisons. The TOWRE is individually administered; Sight Word Reading Efficiency measures students' ability to read real words out of context quickly and accurately and Phonemic Decoding measures students' ability to read pronounceable non-words quickly and accurately. The internal consistency of these measures exceeds .95.

The Gates MacGinitie Reading Test - Comprehension subtest (GMRT-RC)—(MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000) is a group-administered assessment of reading comprehension requiring participants to read passages of narrative and expository text silently and answer relevant comprehension questions. It is considered to be less dependent on word reading and more on oral language proficiency than other similar standardized reading tests (Cutting & Scarborough, 2006). Internal consistency reliability ranges from .91 to .93. GMRT-RC lexile scores were used in analyses as a measure of reading comprehension to provide an objective measure of reading comprehension on a continuous or developmental scale of reading ability that takes word frequency and syntactic complexity into account (Stenner, Burdick, Sanford, & Burdick, 2006). Lexile scores are preferred to other types of reading comprehension scores because they allow for the use of an equal interval scale that is not age-adjusted and more revealing of age-related factors should they exist. Note that grade and age are controlled in the analyses.

Proactive Verbal Interference Task (Pimperton & Nation, 2010)—The proactive verbal interference task is an individually administered computerized task that measures verbal suppression (on interference trials) and verbal cued recall (on non-interference trials). The task includes 4 practice trials and 24 test trials, with all trials consisting of either a single or double block structure.

Each trial began with a visual prompt 'Ready?' followed directly by an audible list of four stimulus words. In the single block trial, a list of words was followed by the visual presentation of a question mark (?). In the double block trial, the list of words was followed by the presentation of an 'X' on the screen, and was followed by an audible list of four additional words, and then a questions mark (?). The 'X' prompted students to forget the first list of four words and focus on remembering the second list of four words. Single block trials are presented to ensure that students pay attention to the first list they hear because they are required to recall from the first list on a third of the trials (i.e., on single block

trials). To prevent rehearsal, after completion of the final word list and the presentation of the question mark, students were required to shadow a list of 20 numbers (a longer shadowed list than in Pimperton & Nation, 2010), presented verbally by the examiner. Immediately after shadowing, students were asked to recall a word from the list in response to a category cue (e.g. Can you remember the word that was a type of pet?).

Half of the 16 double block trials consisted of ‘interference’ trials; half consisted of ‘no-interference’ trials. In the interference trials both the first block (e.g., with the foil word ‘cat’) and the second block (with the target word ‘dog’) contained a category cue matching word. In the no-interference trials, a word matching the category cue (e.g. type of pet) was only presented in the second block (e.g., with the target word ‘dog’). Students’ responses were scored as correct if they accurately produced the target word. A measure of suppression was provided by the number of interference trials correctly responded to out of 8 (Pimperton & Nation, 2010). A measure of verbal cued recall was provided by computing the number correct out of the 8 non-interference trials. Reliability coefficients for the proactive interference task (Kuder-Richardson 20 [KR-20]) for the entire sample were .63 for raw and standardized scores. Because reliability of proactive interference measures in group-based studies of comprehension is typically not reported we cannot compare reliability of this measure in our study with reliability for similar measures used in those other studies. However, this value is consistent with studies in which reliability of proactive interference paradigms is reported (e.g., Friedman & Miyake, 2004).

Results

Overview of Analyses

A set of hierarchical regression models were run to allow estimation of the relation between reading comprehension and the proactive interference variables (recall on interference trials and on no-interference trials) while controlling for other related characteristics. For all models, the variable modeled was the GMRT-RC lexile score. The baseline model included age in years at the time the Gates was administered, Grade in school and KBIT-2 Matrices. In the next model, word reading predictors were added. These included TOWRE-PD, TOWRE-SW and WJIII-LWID. Nonverbal cognitive ability (KBIT-2 Matrices) and the word reading predictors were in the models to parallel as closely as possible, in a regression-based design, the variables that are typically “equated” in groups of good and poor comprehenders who are then compared on suppression (e.g., Pimperton & Nation, 2010). In the next model, either the No-interference or the Interference score was entered and the increment in variance accounted for tested. In the next model, both No interference and Interference scores were entered simultaneously. The no-interference score was included in these models so that an effect for the interference trials could be uniquely ascribed to the suppression of irrelevant information rather than to general cued recall abilities. As Friedman and Miyake (2004) point out, “task impurity” is a considerable issue in measurement of executive functions in general, and of inhibition functions, in particular. The use of the no-interference score as a control variable in the models was meant to control for general abilities (cued recall) that are required on interference trials, but which do not assess suppression. Finally, models that included the interaction of grade and the interference or no-interference score

were run as well as models that included a vector to allow for a quadratic relation between the interference measures and reading comprehension in order to test the linearity of the relation of suppression and reading comprehension across the distribution of comprehension scores. The entire set of models was then run only on participants whose KBIT-2 Matrix score was greater than or equal to 90 (25th percentile), so that our sample was nearly as high functioning as the sample reported on in Pimperton and Nation (2010).

Case-Control comparisons were made via *t*-tests where the matched more skilled and less skilled comprehenders were compared on No interference trial performance and on Interference trial performance. This was done both on the sample that included 48 students per group and also on the sample that remained when students with KBIT-2 Matrices lower than 90 were deleted. In these comparisons, 31 students remained in the less skilled comprehender group and 33 remained in the more skilled comprehender group.

Regression Analyses

The results of the regression analyses are presented separately for the entire sample and for the subset where performance on the KBIT-2 Matrices subset was at or above the 25th percentile. Scores for both the larger sample and the subset are in Table 1. In the set of analyses on the entire sample ($N=766$) where the proactive interference scores were used to predict GM Passage comprehension Lexile score, the base model (age, grade and KBIT-2-Matrices) was significant, ($F_{(3,762)} = 110.06, p < .0001$) accounting for 30.24% of the variance in comprehension. When the word reading variables were added to the model, the variance accounted for rose to 39.34% and the increment was significant ($F_{(3,759)} = 38.01, p < .0001$). The addition of the No interference score increased variance accounted for by 1.9% ($t_{(758)} = 4.91, p < .0001$). Neither the No interference score by grade interaction nor the quadratic effect of No interference added significantly to this model. The addition of the Interference score to the model containing both the base variables and the word reading variables resulted in a significant increase in variance accounted for of .9% ($t_{(758)} = 3.39, p < .0001$). When both Interference and No Interference scores were added simultaneously, they jointly accounted for 2.19% of the variance in reading comprehension ($F_{(2,757)} = 14.22, p < .0001$). Neither the Interference trials by grade interaction nor the quadratic effect of Interference trials added significantly to this model.

In the parallel set of analyses where only participants with KBIT-2 Matrix standard scores greater than or equal to 90 were included ($N = 478$), the base model accounted for 31.37% of the variance in reading comprehension ($F_{(3,474)} = 72.22, p < .0001$). Addition of the word reading variables raised the variance accounted for to 40.37% and the increment was significant ($F_{(3,471)} = 23.71, p < .0001$). The addition of the No interference score resulted in accounting for an additional 1.48% of the variance which was significant ($t_{(470)} = 3.46, p = .0006$). The addition of the Interference score to the model containing both the base variables and the word reading variables resulted in a significant increase in variance accounted for of .9% ($t_{(470)} = 2.57, p < .01$). When both Interference and No Interference scores were added simultaneously, they jointly accounted for 1.79% of the variance in reading comprehension ($F_{(2,469)} = 7.25, p = .0008$). Neither the Interference trials by grade interaction nor the quadratic effect of Interference trials added significantly to this model.

Matched group analyses

From the larger group we matched more and less skilled comprehenders as closely as possible for age and grade and then followed the general procedures outlined in Pimperton and Nation (2010) as follows. The less skilled comprehender group had scores below the 25th percentile for their grade on the GMRT-RC. Less and more skilled comprehenders were matched for their scores on the KBIT-2 Matrices test. Less skilled comprehenders had average or above average nonword reading assessed by the Phonemic Decoding subtest on the TOWRE. More skilled comprehenders had reading comprehension scores that were similar to their reading accuracy. Consistent with the approach used in the regression analyses, we also tested for group differences using the subset of participants with scores on the KBIT-2 Matrices at or above the 25th percentile. Scores for the groups on the matching variables, reading comprehension, and recall on the interference and no-interference trials are in Table 2.

We estimated the *t*-tests and calculated effect sizes simply to make comparisons across the results, not because they are appropriate. An independent *t*-test comparing the matched more skilled ($N = 48$) and less skilled ($N = 48$) comprehender groups on their performance on the no-interference trials was not significant ($t_{(94)} = -1.63$, $p = .11$). The effect size d was $-.33$, with the more skilled comprehenders scoring higher than the less skilled comprehenders. The effect size in Pimperton and Nation (2010) was slightly higher: $.54$, but the comparison between their skilled and less skilled comprehenders was not statistically significant for the no-interference trials. An independent *t*-test comparing the matched more and less skilled comprehender groups on their performance on interference trials was not significant ($t_{(94)} = -1.54$, $p = .13$). The effect size d was $-.31$, with the skilled comprehenders scoring higher than the less skilled comprehenders. The effect size in the Pimperton and Nation (2010) study was much larger: -1.10 , and the difference between the skilled and less skilled comprehenders was significant for the interference trials in their study. When only participants with KBIT-2 Matrices scores at or above the 25th percentile were included in the independent *t*-tests, none of the results were significant.

Discussion

Informing Reading for Understanding through Basic Research Studies in Adolescents' Reading Comprehension

In the RFU project, Promoting Adolescents' Comprehension of Text (PACT), both intervention and basic research studies, like the one presented here have been conducted. The basic research studies are meant to inform models of comprehension and provide information relevant to assessment and intervention for adolescent readers. It is relevant for models of comprehension and to educational practice to understand which of the many proposed comprehension-related variables in the literature are *pressure points* for comprehension (Perfetti & Adlof, 2012); that is which comprehension-related abilities are integral and importantly related to comprehension. To generate knowledge in basic research studies about reading for understanding in adolescents, we have modeled hypothesized sources of direct and indirect influences on reading comprehension (Ahmed et al., 2015; Barnes et al., 2015; Barth et al., 2015; Denton et al., 2015). Based on cognitive process

models of reading comprehension (e.g., Landscape Model, van den Broek et al., 2005 and the Structure Building Framework, Gernsbacher, 1996) we have also considered whether and to what extent several general cognitive abilities - memory, attention, response inhibition, and suppression – are related to comprehension (e.g., Arrington et al., 2014; current study).

Because many studies relating general cognitive abilities to individual differences in comprehension are conducted comparing good and poor comprehenders matched on word reading and nonverbal cognitive abilities, and because of the methodological concerns with such designs, we use regression based approaches in many of our studies to both estimate the *size* of the relation of cognitive correlates to comprehension and determine whether those relations are *invariant* over the distribution of reading comprehension levels. In keeping with the overall purpose of the PACT project, the questions we addressed in this study were: 1) is there a relation of suppression and reading comprehension; 2) is that relation invariant across the distribution of reading comprehension; 3) what is the size of this relation; and 4) are findings from regression-based analyses similar to those using a more typical extreme groups comparisons. The answers to the first three questions are important for determining whether suppression is an important target for assessment and intervention in reading for understanding, and also for informing comprehension models that propose a particular role for suppression in the construction of text representations. To our knowledge, the 4th question represents the first direct comparison within the same sample, of an approach using both regression-based and extreme group designs, which addresses an important methods issue in understanding individual differences in reading for understanding.

The Relation of Suppression and Reading Comprehension

General-purpose cognitive processes such as the ability to focus on task-relevant and suppress task-irrelevant information in working memory have been hypothesized to partly account for individual differences in many aspects of higher cognition including reading comprehension. Models of comprehension, such as the Structure Building Framework (Gernsbacher, 1996) explicitly include suppression as a primary cognitive mechanism that allows for construction of an accurate representation of text. According to this model, an accurate and accessible representation of the text is constructed through the maintenance of contextually-relevant information and the suppression of irrelevant information (e.g., Cain, 2006; Gernsbacher, 1996). Supporting this are several studies in which differences on tasks tapping suppression have been found for good and poor comprehenders matched for decoding ability and, in some studies, also matching for nonverbal cognitive ability (Barnes et al., 2004; Borella et al., 2010; Caretti et al., 2005; Gernsbacher & Faust, 1991; Pimperton & Nation, 2010). However, studies using extreme group designs are not able to address whether findings at the extremes also represent those in the middle of the distribution, and because such designs are associated with inflated effect sizes and regression to the mean (Preacher et al., 2005), they may not produce accurate estimates of the size of effects. These considerations motivated our regression-based approach to address questions about the relation of suppression and reading comprehension.

In these regression analyses, word decoding efficiency and nonverbal IQ, the variables that are matched between groups of skilled and less skilled comprehenders in extreme group designs were controlled. The ability to recall words when suppression is not required (recall on no-interference trials) was also included in the models in order to control for extraneous task variables not related to suppression per se. Suppression (recall on interference trials) accounted for significant unique variance in reading comprehension. However, the effect was small, about 1%. The effect was invariant across the distribution of reading comprehension, suggesting that the relation of suppression to reading comprehension in adolescent readers is similar regardless of their comprehension level. Using an extreme groups approach in which the groups were matched on age, grade, word reading and nonverbal cognitive variables, no difference in suppression was found for a group of 48 less skilled comprehenders compared to 48 more skilled comprehenders, despite the use of larger groups than those typically employed in group-based designs (but see Caretti et al., 2005).

The findings from the regression analyses support the hypothesis that suppression is uniquely predictive of reading comprehension (Arrington et al., 2014; Pimperton & Nation, 2010) and contrast with those from Christopher et al. (2012) who showed that once other cognitive abilities (e.g., processing speed and working memory) were in the same model as their inhibition construct, inhibition had no unique predictive ability for reading comprehension.

There are two potential explanations for the different findings concerning the relation of inhibition and comprehension. First, as suggested by Pimperton and Nation (2010), the relation of suppression to reading comprehension may be domain specific; that is, suppression in the language domain is related to reading comprehension. Unlike Christopher et al. (2012), the current study used a verbal test of suppression. Another possibility has to do with the particular type of inhibitory function assessed in the two studies. Inhibition is a *family* of functions (Friedman & Miyake, 2004), each of which could be related in different ways to different aspects of reading (see Arrington et al., 2014). The type of inhibition measured in the Christopher et al. (2012) study was response inhibition whereas the type of inhibition measured in the current study was resistance to proactive interference (i.e., suppression). A study employing an extreme group design also found a difference between good and poor comprehenders on proactive interference tasks, but not on tasks assessing response inhibition (Borella et al., 2010).

One advantage of the regression approach is that it allowed us to test whether the relation of suppression and comprehension is invariant across the distribution of reading comprehension performance. Extreme groups methods assume or impose a linear relation between two variables, but if the relation of two variables is not actually linear, then extreme group designs will lead to model misspecification (Jackson & Butterfield, 1989; Preacher et al., 2005). With the regression approach, in contrast, it is possible to estimate whether the form of the relation is linear or whether higher order functions are needed to capture the relation. In this study, there was no evidence that the relation of suppression and comprehension differed as a function of comprehension level, suggesting that the mechanism by which suppression accounts for variance in performance for good and poor comprehenders is similar to that for the larger distribution of adolescent comprehenders.

It is worth noting that in the present study, a “significant” difference in suppression was not obtained between skilled and less skilled comprehenders using an extreme groups approach. In the extreme groups approach, the effect estimated is the size of the difference between groups on the cognitive variable, suppression, and this will depend on the strength of the association between comprehension and suppression and also on the levels of comprehension chosen to define the groups. The difference in comprehension can be fairly arbitrary depending on the selection of the sample. In contrast, the regression approach uses the supposed antecedent/cognitive variables to predict reading comprehension. Causality goes from the components to the achievement variable and the relation estimated is between comprehension and that part of the cognitive variables that are unrelated to the covariates (Jackson & Butterfield, 1989). We would argue that because both extreme group and regression analyses are a form of correlation (neither is a true experiment even though both assume an underlying causal model that drives matching or covarying), using the distribution of scores is more likely to result in a more stable and less arbitrary estimate of relations of the variables of interest in the underlying causal model than will trying to select small subgroups from a larger sample. It is interesting to think about the extent to which group selection using particular matching variables in these studies may result in misrepresentative samples. In Pimperton and Nation (2010), 25% of the larger screened sample was able to be matched to form the skilled and less skilled comprehenders groups. In the current study, 13% of the larger sample was able to be matched.

What are we to make of the lack of significant group difference in suppression using an extreme groups design? As discussed above, we suggest that finding group differences in such designs is somewhat arbitrary and based on factors often unrelated to the main questions of interest. While it might seem more convincing if we had observed the same effects in the present study as have been found in other studies of extreme groups, one limitation of extreme group studies is the presence of selection bias brought about by the formation of the groups. If all such studies drew their samples from the same populations in the same ways, then these biases would be expected to operate in a similar way in all studies. Because studies differ systematically from one another, these selection biases will also differ across studies and affect the replicability of findings. Thus, the regression method is not superior because it replicates the findings of the extreme groups analyses; the regression method is superior because it statistically controls selection biases that affect comparisons of good and poor comprehenders. Furthermore, we note that studies that do not find effects are less likely to be published. Of course, it is not possible to determine whether this is the case for the relation of suppression and reading comprehension. However, the advantage of the current study lies in the comparison of findings from regressions and those from the extreme group design where the groups are drawn from the same population of students. Although the group based analysis did not result in a significant difference on suppression, the regressions produced a significant, albeit small effect, as discussed next.

Estimating the Size of the Suppression-Comprehension Relation

In thinking about the importance of cognitive processes for reading comprehension, it is necessary to ask not only whether a particular process is related to comprehension, but also what is the size of that relation. In order to accurately estimate the size of an effect

regression provides a more precise estimate than extreme group designs when the sample size is adequate. Because regression uses the continuum of suppression and comprehension scores and does not apply arbitrary cut-offs, estimates of effect size are less likely to be inflated. Because matching reduces the representativeness of the samples (in an attempt to equalize samples on a set of variables when those samples are actually drawn from populations that are intrinsically non-equivalent), the external validity of the study is decreased and the magnitude of regression effects is increased (Jackson & Butterfield, 1989).

In the regression analyses, the effect of suppression on reading comprehension was significant, but it was small. Although the findings support the hypothesis that suppression abilities for language-based materials are uniquely related to reading comprehension, the small size of this effect is important for thinking about what have been termed “pressure points” of comprehension (Perfetti & Adlof, 2012). A pressure point refers to a component of comprehension that is intrinsic to comprehension as opposed to simply a cognitive correlate of comprehension. Why do we care about pressure points? Something that is intrinsic to comprehension is likely to be relevant for comprehension assessment and instruction and, depending on developmental timing of comprehension-related skills, for the prevention of reading comprehension difficulties as well. Skills that *consistently* show *substantial* variation between individuals who differ in comprehension ability are considered to be good candidates for “pressure points” (Perfetti & Adlof, 2012). The current findings from the group-based analyses suggest that suppression does not *consistently* vary between good and poor comprehenders. The findings from the regression analyses suggest that although suppression and comprehension are related, this relation is not sufficiently substantial or of such a magnitude that we might consider it to be of fundamental importance to our assessment or instruction of comprehension.

In comparing the estimates of effects across the Pimperton and Nation (2010) study and our two studies, predictable patterns emerge. The largest effects of $d = -1.10$ and $d = -.54$ (which we calculated based on means and SDs reported by Pimperton and Nation (2010) were those based on the smallest sample size ($n = 28$). Knowledge of the effect of sample size on sampling error would lead us to expect the largest variability in effects in studies with small sample sizes, even when these effects come from experimental designs with participants randomly assigned to groups. The effect of creating groups by selecting homogeneous sets of participants with extreme scores exaggerates this effect. In contrast, the effects we found in our group study with a larger sample size ($n = 96$) $d = -.31$ and $d = -.33$ are considered between medium and small. With larger sample sizes we expect less variability from study to study, but we expect that there is still effect size inflation due to the selection of extreme and homogeneous groups. In our regression analyses, the variance accounted for by interference trials was less than 1%, which is a small effect. It is based on a very large sample ($n = 766$) so it is expected to have small sampling variance and it is also not inflated by any manipulation of the sample.

In addition to sample size differences we note some other ways in which our study differs from other studies in which suppression has been assessed in extreme group designs. These factors could also account for differences in findings between studies. Our participants were

older than those in most skilled/less skill comprehender group studies, including Pimperton and Nation (2010). Although ceiling effects on the recall task for older students did not appear to be an issue, developmental differences in the relation of inhibitory processing to reading comprehension could be important. Typically developing younger children may have more difficulty ignoring irrelevant information during reading than do older typical readers (Pike, Barnes, and Barron, 2010). On the other hand, there is evidence for a relation of suppression and reading comprehension in ambiguity resolution studies with adults (e.g., Gernsbacher & Faust, 1991); however, Henderson, Snowling, & Clarke (2013) suggest that findings for poor comprehenders in these paradigms are actually due to less familiarity with the lower frequency meanings of ambiguous words, a vocabulary-based explanation rather than a suppression-based interpretation. Whether the differences between findings in the current study and those with younger children using proactive interference paradigms has to do with moderation of the relation of suppression and comprehension by age cannot be determined from the current findings. The development of inhibitory processing, for example, is complex, may be task-dependent (Friedman & Miyake, 2004), and interact with other cognitive processes such as memory demands (Davidson, Amso, Anderson, & Diamond, 2006). Furthermore, the factors that most affect comprehension may vary somewhat across development (e.g., Tighe & Schatschneider, 2014).

We are also cognizant that the level of cognitive ability in other extreme groups comprehender studies is higher than that in the current study. In an attempt to address this issue, we ran the regression and group analyses only with those individuals whose nonverbal cognitive level was at or above the 25th percentile. This, however, did not change the pattern of findings or the estimate of effects. Regardless, the findings from the regression analyses suggest that in a large and representative group of adolescents, the effect of suppression on reading comprehension is significant, but small.

Limitations

There are three main limitations in the current study. One has to do with the capacity to generalize to the broader population of adolescents. Although we have claimed that the use of the entire distribution of comprehenders in the current study addresses some of the methodological issues related to generalization and the accurate estimation of effects, it is important to note that our sampling strategy (oversampling for less skilled comprehenders and screening out students with decoding accuracy below the 20th percentile) resulted in a sample in which word reading abilities were truncated. Therefore, we cannot say with certainty that the relation of suppression and comprehension found in this study generalizes to the entire distribution of adolescent readers.

Second, there are issues related to measurement in terms of how well the tasks used in the current study capture, at the latent level, the variables of interest – suppression and reading comprehension. In matched group designs, construct validity can be weak because participants are matched on observed variables. Due to imperfect measurement they are almost certainly not matched on the constructs of interest. However, this is also a weakness of the current regression approach using single indicators as it is highly unlikely that the covariates or the predictor and outcome variables are measured without error. Unless these

variables are measured without error (as is assumed by the regression model) the estimated effects can be biased. It would be of interest to test the relation of suppression and reading comprehension using latent variables. These studies might ensure that a latent variable for verbal suppression be compared to that for nonverbal suppression and that a latent variable assessing resistance to proactive interference be compared to latent variables assessing other types of inhibitory functions. Third, the regression models in the current study were necessarily simplistic because we wanted to conduct a study that was as similar as possible to studies using extreme group designs relating suppression to comprehension. In keeping with this logic, we included only those variables that have been used to match groups in extreme group designs. A better understanding of how suppression and reading comprehension are related, for example, whether suppression mediates the relation of working memory to comprehension (McVay & Kane, 2012), or whether suppression is significant in the presence of general processing speed (e.g., Christopher et al., 2012), requires a more comprehensive model-driven set of analyses than those used in the current study.

Conclusions

It is clearly important for both theory and educational practice to obtain accurate estimates of the size of the relation of reading comprehension and cognitive variables such as suppression. In this instance, the regression-based design shows a small unique effect of suppression on reading comprehension. The failure to find a significant suppression effect in the extreme group design may simply reflect the measurement issues that emerge across samples when participants are selected using a variety of covariates and when samples and sample sizes vary across studies. As the present study shows, these problems are less apparent with regression-based designs.

Importantly, the findings suggest that while suppression is uniquely related to comprehension, the size of that effect is not sufficiently large enough to motivate the inclusion of measures of proactive interference in assessments of reading for understanding. Furthermore, because the size of the effect is both small and invariant across comprehension level, the findings also provide no evidence in support of the creation of interventions designed to improve the suppression of irrelevant information during reading for adolescents with low reading comprehension (see Borrella et al., 2010 for a differing view). Although the findings seem to also question a central role for suppression in models of reading comprehension, there is an important caveat with respect to implications for theory. Cognitive models of comprehension, such as the Structure Building Framework, are process models, meaning that they describe the on-line processes involved in constructing text representations. Studies like the one reported here treat particular measures (i.e., intrusion errors in proactive interference paradigm) as a proxy for cognitive “processes” (e.g., suppression). However, there may be important distinctions between how inhibitory processing takes place in real time during reading and the measurement of inhibitory functions as tapped by memory tasks that are administered outside of reading for understanding.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to the University of Texas-Austin as part of the Reading for Understanding Research Initiative and by Award Number P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to the University of Houston. The opinions expressed are those of the authors and do not necessarily represent views of the Institute or the U.S. Department of Education or official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health. The authors wish to acknowledge the invaluable contributions of the middle and high school staff and students in Channelview ISD, Dickinson ISD, Galveston ISD and Humble ISD in Texas. We also thank Hannah Pimperton and Kate Nation for sharing their materials with us.

References

- Arrington CN, Kulesz PA, Francis DJ, Fletcher JM, Barnes MA. The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*. 2014; 18(5):325–346.
- Baddeley A. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*. 2000; 4(11):417–423. [PubMed: 11058819]
- Barnes MA, Ahmed Y, Barth A, Francis DJ. The Relation of Knowledge-Text Integration Processes and Reading Comprehension in 7th-to 12th-Grade Students. *Scientific Studies of Reading*. 2015:1–20. ahead-of-print.
- Barnes MA, Faulkner H, Wilkinson M, Dennis M. Meaning construction and integration in children with hydrocephalus. *Brain and Language*. 2004; 89(1):47–56. [PubMed: 15010236]
- Barth AE, Barnes M, Francis D, Vaughn S, York M. Inferential processing among adequate and struggling adolescent comprehenders and relations to reading comprehension. *Reading and Writing*. 2015; 28(5):587–609. [PubMed: 26166946]
- Borella E, Carretti B, Pelegrina S. The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*. 2010; 43(6):541–552. [PubMed: 20606207]
- Branum-Martin L, Fletcher JM, Stuebing KK. Classification and identification of reading and math disabilities: The special case of comorbidity. *Journal of Learning Disabilities*. 2013; 12:906–915.
- Cain K. Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory*. 2006; 14(5):553–569. [PubMed: 16754241]
- Cain K, Oakhill J, Bryant P. Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing*. 2000; 12(1-2):31–40.
- Campbell, DT.; Kenny, DA. A primer on regression artifacts. Guilford Press; New York, NY: 1999.
- Carretti B, Cornoldi C, De Beni R, Romanò M. Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology*. 2005; 91(1):45–66. [PubMed: 15814095]
- Christopher ME, Miyake A, Keenan JM, Pennington B, DeFries JC, Wadsworth SJ, Olson RK. Predicting word reading and comprehension with executive function and speed measures across development: a latent variable analysis. *Journal of Experimental Psychology: General*. 2012; 141(3):470–487. [PubMed: 22352396]
- Cirino PT, Romain MA, Barth AE, Tolar TD, Fletcher JM, Vaughn S. Reading skill components and impairments in middle school struggling readers. *Reading and Writing*. 2013; 26(7):1059–1086. [PubMed: 24000271]
- Cutting LE, Scarborough HS. Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*. 2006; 10(3):277–299.
- Davidson MC, Amso D, Anderson LC, Diamond A. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*. 2006; 44(11):2037–2078. [PubMed: 16580701]
- De Beni R, Palladino P. Intrusion errors in working memory tasks. Are they related to reading comprehension ability? *Learning and Individual Differences*. 2000; 12:131–143.

- Denton CA, Enos M, York MJ, Francis DJ, Barnes MA, Kulesz PA, Fletcher JM, Carter S. Text-Processing Differences in Adolescent Adequate and Poor Comprehenders Reading Accessible and Challenging Narrative and Informational Text. *Reading Research Quarterly*. 2015; 50:393–416.
- Ellis AW. The cognitive neuropsychology of developmental (and acquired) dyslexia: A critical survey. *Cognitive Neuropsychology*. 1984; 2:169–205.
- Engle RW, Kane MJ. Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation*. 2004; 44:145–200.
- Friedman NP, Miyake A. The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General*. 2004; 133(1):101–135. [PubMed: 14979754]
- Gernsbacher, MA. The structure – building framework: What it is, what it might also be, and why. In: Gerrig, RJ., editor. *Models of Understanding Text*. Lawrence Erlbaum Associates; Mahwah, NJ: 1996. p. 289–311.
- Gernsbacher MA, Faust ME. The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1991; 17(2):245–262.
- Henderson L, Snowling M, Clarke P. Accessing, integrating, and inhibiting word meaning in poor comprehenders. *Scientific Studies of Reading*. 2013; 17(3):177–198.
- Jackson NE, Butterfield EC. Reading-level-match designs: Myths and realities. *Journal of Literacy Research*. 1989; 21(4):387–412.
- Just MA, Carpenter PA. A capacity theory of comprehension: individual differences in working memory. *Psychological Review*. 1992; 99(1):122. [PubMed: 1546114]
- Kaufman, AS.; Kaufman, NL. Kaufman Brief Intelligence Test, Second Edition (KBIT-2). American Guidance Service; Circle Pines, MN: 2004.
- Lee K, Ng EL, Ng SF. The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*. 2009; 101(2):373.
- MacGinitie, WH.; MacGinitie, RK.; Maria, K.; Dreyer, LG.; Hughes, KE. Gates-MacGinitie reading tests, fourth edition. Riverside; Itasca, IL: 2000.
- McVay JC, Kane MJ. Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology*. 2012; 141(2):302. [PubMed: 21875246]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*. 2000; 41(1):49–100. [PubMed: 10945922]
- Miyake, A.; Shah, P., editors. *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press; Cambridge, United Kingdom: 1999.
- Palladino P, Cornoldi C, De Beni R, Pazzaglia F. Working memory and updating processes in reading comprehension. *Memory & Cognition*. 2001; 29(2):344–354. [PubMed: 11352218]
- Perfetti, CA.; Adlof, SM. Reading comprehension: A conceptual framework from word meaning to text meaning. In: Sabatini, JP.; Albro, ER.; O'Reilly, T., editors. *Measuring up: Advances in how we assess reading ability*. Rowland & Littlefield; Lanham, Maryland: 2012. p. 3–20.
- Pike MM, Barnes MA, Barron RW. The role of illustrations in children's inferential comprehension. *Journal of Experimental Child Psychology*. 2010; 105(3):243–255. [PubMed: 19954799]
- Pimperton H, Nation K. Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language*. 2010; 62(4):380–391.
- Preacher KJ, Rucker DD, MacCallum RC, Nicewander WA. Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychological Methods*. 2005; 10(2):178. [PubMed: 15998176]
- Stenner AJ, Burdick H, Sanford EE, Burdick DS. How accurate are lexile text measures? *Journal of Applied Measurement*. 2006; 7(3):307. [PubMed: 16807496]

- Tighe EL, Schatschneider C. A dominance analysis approach to determining predictor importance in third, seventh, and tenth grade reading comprehension skills. *Reading and Writing*. 2014; 27(1): 101–127. [PubMed: 26346315]
- Torgesen, JK.; Wagner, R.; Rashotte, C. TOWRE–2 test of word reading efficiency. Pro-Ed; Austin, TX: 1999.
- van den Broek P, Rapp DN, Kendeou P. Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*. 2005; 39:299–316.
- Wilson SP, Kipp K. The development of efficient inhibition: Evidence from directed-forgetting tasks. *Developmental Review*. 1998; 18(1):86–123.
- Woodcock, RW.; McGrew, K.; Mather, N. Woodcock-Johnson battery. 3rd. Riverside Publishing; Itasca, IL: 2007. Normative Update

Table 1

Means (Standard Deviations) for Measures in Regression Analyses

Measure	All participants in sample (<i>n</i> = 766)	Participants with KBIT Matrices at or above the 25 th percentile (<i>n</i> = 478)
WJ-III LWID	99.8 (8.4)	101.5 (8.9)
TOWRE-SW	93.5 (10.1)	94.5 (10.2)
TOWRE-PD	96.4 (11.9)	97.4 (12.02)
GMRT-RC	92.5 (12.7)	94.3 (13.0)
KBIT-2 Matrices	93.5 (10.7)	99.9 (7.7)
PV Interference	4.3 (1.7)	4.4 (1.6)
PV No Interference	5.7 (1.5)	5.8 (1.5)

Note. All means represent standard scores with the exception of the scores for interference and no-inferences trials (x/8). WJ-III LWID = Woodcock Johnson-III Tests of Achievement, Letter Word Identification subtest; TOWRE-SE = Test of Word Reading Efficiency, Sight Word Reading Efficiency subtest; TOWRE-PD = Test of Word Reading Efficiency, Phonemic Decoding subtest; GMRT-RC = Gates MacGinitie Reading Test, Comprehension subtest; KBIT-2 Matrices = Kaufman Brief Intelligence Test – 2, Matrices subtest; PV Interference = Proactive Verbal Interference Task, interference trials; No Interference = Proactive Verbal Interference Task, no-interference trials.

Table 2

Means (Standard Deviations) for Measures in Matched Case Analysis

Measure	Less skilled comprehenders (<i>n</i> = 48)	More skilled comprehenders (<i>n</i> = 48)
WJ-III LWID	98.0 (7.8)	99.7 (7.1)
TOWRE-SW	98.3 (9.3)	98.6 (8.9)
TOWRE-PD	102.0 (10.0)	99.6 (10.1)
GMRT-RC	78.0 (7.0)	104.9 (8.3)
KBIT-2 Matrices	94.4 (11.7)	96.8 (11.4)
PV Interference	4.1 (1.6)	4.6 (1.5)
PV No Interference	5.4 (1.7)	5.9 (1.6)

Note. All means represent standard scores with the exception of the scores for interference and no-inferences trials (x/8). WJ-III LWID = Woodcock Johnson-III Tests of Achievement, Letter Word Identification subtest; TOWRE-SE = Test of Word Reading Efficiency, Sight Word Reading Efficiency subtest; TOWRE-PD = Test of Word Reading Efficiency, Phonemic Decoding subtest; GMRT-RC = Gates MacGinitie Reading Test, Comprehension subtest; KBIT-2 Matrices = Kaufman Brief Intelligence Test – 2, Matrices subtest; PV Interference = Proactive Verbal Interference Task, interference trials; PV No Interference = Proactive Verbal Interference Task, no-interference trial.