From Words to Text: Inference making Mediates the Role of Vocabulary in Children's Reading Comprehension

Hanne Trebbien Daugaard (corresponding author)

University of Copenhagen

trebbien@hum.ku.dk

Kate Cain

Lancaster University

k.cain@lancaster.ac.uk

Carsten Elbro

University of Copenhagen

ce@hum.ku.dk

Paper accepted for publication in Reading and Writing, 25 May 2017

Acknowledgements: This study was undertaken as part of the first author's PhD project, which was embedded in the project *Language and Cognition – Perspectives from Impairment* (LaCPI) 2011-2014. LaCPI was supported by a grant to professor Elisabeth Engberg-Pedersen from The Danish Council for Independent Research | Humanities, by The Faculty of Humanities, University of Copenhagen, and by The Department of Nordic Studies and Linguistics, University of Copenhagen. The members of LaCPI also included postdoc Rikke Vang Christensen and eighteen students. We

would like to thank all the members of LaCPI as well as the children, parents, and schools who participated in the study.

Abstract

We examined the relationship between inference making, vocabulary knowledge, and verbal working memory on children's reading comprehension in 62 6th graders (aged 12). The effect of vocabulary knowledge on reading comprehension was predicted to be partly mediated by inference making for two reasons: Inference making often taps the semantic relations among words, and the precise word meanings in texts are selected by readers on the basis of context. All independent variables were significantly and moderately correlated with reading comprehension. In support of our prediction, the link between vocabulary knowledge and reading comprehension was significantly mediated by inference making even when verbal working memory was controlled. An alternative mediation hypothesis (vocabulary as a mediator of the effect of inference making on comprehension) was not supported by the data. The study replicates and extends the findings of earlier work (Cromley & Azevedo, 2007; Segers & Verhoeven, 2016; Ahmed et al., 2016).

From Words to Text: Inference making Mediates the Role of Vocabulary in Children's Reading Comprehension

Vocabulary is crucial to text comprehension. Knowledge of the majority of the words in a text is imperative for understanding what the text is about (Nation, 2009). There is empirical support for this claim. Vocabulary predicts children's reading comprehension both concurrently and longitudinally (Nation & Snowling, 2004) and evidence for a reciprocal relationship between vocabulary development and reading comprehension development has been found (Verhoeven, van Leeuwe, & Vermeer, 2011). Training studies also support the importance of vocabulary for comprehension. Primarily, vocabulary instruction can improve children's reading comprehension when comprehension is assessed with custom measures that include the taught words (Elleman, Lindo, Morphy, & Compton, 2009). Few studies have also found a small, but significant, effect on standardized reading comprehension measures (e.g., Beck, McKeown, Omanson, & Perfetti, 1982; Vadasy, Sanders, & Logan Herrera, 2015).

Inference making is another important component of text comprehension (Graesser, Singer, & Trabasso, 1994). In order to construct a meaningful and coherent representation of the affairs described in the text (a mental model or situation model) (Johnson-Laird, 1983; Kintsch, 1998) the reader makes connections between ideas described in the text and his or her relevant background knowledge to generate information (inferences) that is not explicitly stated in the text (van den Broek, Risden, & Husebye-Hartmann, 1995). Many of these inferences can be characterized as establishing either referential coherence (connecting objects, characters and other entities) or causal coherence (connecting events) (van den Broek, White, Kendeou, & Carlson, 2009). Like vocabulary, inference making predicts variation in children's reading comprehension both concurrently and longitudinally (Oakhill & Cain, 2012) and inference training can improve children's reading comprehension (e.g., Elbro & Buch-Iversen, 2013).

In the present study, we examined the relations between vocabulary, inference making, and reading comprehension in 6th graders. Additionally, we examined the role of working memory in supporting these skills. Multi-component models of reading comprehension allow for interactions

between lower level skills (e.g., vocabulary knowledge and retrieval of word meanings) and higher level skills (e.g., inference making) in the process of comprehension (e.g., Perfetti, Landi, & Oakhill, 2005). But the contributions of vocabulary and inference making to individual differences in children's reading comprehension have often been considered separately, since the goal of many empirical studies have been to confirm the significance of a specific component for reading comprehension (e.g., Nation, Clarke, Marshall, & Durand, 2004; Cain & Oakhill, 1999). However, evidence for a complex relationship between vocabulary and inference making in children's reading comprehension has emerged from recent empirical findings. Vocabulary is an important predictor of children's inference making (Cain & Oakhill, 2014; Currie & Cain, 2015) and evidence of an indirect (i.e., mediated) effect of vocabulary on reading comprehension through inference making has been found (Cromley & Azevedo, 2007). Cromley and Azevedo tested the Direct and Inferential Mediation (DIME) model of reading comprehension with high school students and found that vocabulary had a moderate direct effect on reading comprehension and also a small but significant indirect effect through inference making. Ahmed et al. (2016) conducted a replication study on the DIME model of Cromley and Azevedo (2007) with a large sample of middle and high school students (grades 7 to 12). The indirect effect of vocabulary was not significant at any grade level when the model was replicated with observed variables. However, it was significant on grade level 9-12 when the model was replicated with latent variables and also on grade level 8 when a general factor was taken into account. A recent study by Segers and Verhoeven (2016) also supports an indirect effect of vocabulary on reading comprehension through inference making, finding that measures of syllogistic reasoning mediated the effect of vocabulary on children's text comprehension (grade 4). In addition, support comes from two studies with adults: a replication of the DIME model with college students (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010) and a think-aloud study with university students (Cromley & Wills, 2016).

Why would inference making mediate the effect of vocabulary on reading comprehension? Cromley and Azevedo (2007) ascribed the indirect effect of vocabulary to reading comprehension through inference making to the fact that whether the reader knows a certain word or not can be

crucial for inference making. But at least two more reasons for expecting a mediation effect should be considered.

First, inference making might tap knowledge about semantic relations *between* words. In Figure 1, the first example, the participant is required to make an inference (weapons and food were loaded in the naval port), which establishes referential coherence (cf. van den Broek et al., 2009). Appreciating that *facility* is superordinate to *naval port* may serve as a knowledge foundation for this inference. In the second example, the participant is required to make an inference (water from the Nile made the bread rise), which establishes causal coherence. Here, recognizing the relation between the words *yeast*, *bake* and *bread* (the baking of bread typically involves yeast) and the contrast of typical (risen) *bread* and (flat) *crispbread* provide a knowledge foundation for inferring that water from the Nile made the bread rise.

Second, most words have several meanings (related or unrelated), and the relevant and precise meaning of a word in a text therefore has to be selected as appropriate to the context i.e., the surrounding text (Miyake, Just, & Carpenter, 1994). Consider for instance the meaning of *facility* in the following sentences: "*In the facility weapons and food were being loaded"* and *"He had a facility for languages.*" Although the same phonological and orthographic form, 'facility' refers to a place for a particular purpose in the first sentence, but an aptitude in the second.

[FIGURE 1]

The results of Cromley and Azevedo (2007), Ahmed et al. (2016), and Segers and Verhoeven (2016) lend support to the proposal that inference making is supported by word knowledge (see also, Cain & Oakhill, 2014; Perfetti, Yang, & Schmalhofer, 2008). Importantly though, neither of the studies assessed whether the data supported an alternative hypothesis: that vocabulary mediates the effect of inference making on reading comprehension. Although vocabulary is an unlikely mediator during reading, vocabulary may attain a mediating role over time. When children read texts they are exposed to rich and varied word use and learning of new words are made possible (Cunningham & Stanovich, 2001). Deriving unknown word meanings from text – by the application of inference skills – is likely to play a role in this process (Nagy, Anderson & Herman, 1987). This relationship between

inference making and word learning could give rise to at mediation effect in cross-sectional data since current vocabulary is an indicator of word learning skills.

A further limitation of the previous studies on the indirect effect of vocabulary through inference making lies in the measures used: in the study by Cromley and Azevedo (2007) vocabulary was assessed as reading vocabulary and, not surprisingly, strongly related to word reading skills (r = .61). In the study by Segers and Verhoeven (2016), vocabulary was also assessed with a pen and pencil test, in which children were asked to choose a synonym (from four alternatives) of an underlined word in a sentence. Interestingly, an important finding by Ahmed et al. (2016) was that when method bias (arising from the reading requirements of component skill test formats) was controlled statistically, this reduced the importance of vocabulary and increased the importance of inference making. The vocabulary variable by Ahmed et al. (2016) included a reading vocabulary test (the same test as Cromley and Azevedo), a word-learning test and an oral vocabulary test (administered with half of the participants).

Inference making was measured in very different ways in the three studies. Cromley and Azevedo (2007) measured inference making with just 8 items from a multiple-choice test. Ahmed et al. (2016) derived several measures from two process-based inference tests. The majority of measures reflected processing speed. Segers and Verhoeven (2016) applied two measures of syllogistic reasoning as proxies of the kind of inference making ability that is important in reading comprehension. However, although syllogistic reasoning and inference ability in text comprehension is significantly correlated, the correlation is weak to moderate (Osana, Lacroix, Tucker, Idan and Jabour, 2007) and the status of logic-based inferences such as syllogisms in reading comprehension is open to discussion (Graesser, Wiemer-Hastings & Wiemer-Hastings, 2001). Clearly, replication with other more comprehensive measures of the constructs – as well as assessment of the alternative hypothesis – is required in the process of confirming whether inference making mediates the effect of vocabulary on comprehension.

The role of working memory. Working memory is a capacity-limited memory system that is engaged in the simultaneous short-term storage and processing of information and interacts with long-

term memory (Baddeley, 2012). Verbal working memory is assumed to be crucial to reading comprehension because the reader has to store and integrate linguistic information continuously while reading (Kintsch & Rawson, 2005) and studies confirm a relation with reading comprehension (e.g., Daneman & Carpenter, 1980; Arrington, Kulesz, Francis, & Fletcher, 2014) as well as with many other complex cognitive skills and behaviours (cf. Baddeley, 1992; Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). In Baddeley's framework of working memory, specialised subsystems undertake temporary storage of phonological and visuo-spatial information (i.e., phonological loop, visuo-spatial sketchpad), as well as multidimensional representations (episodic buffer) whereas simultaneous processing of either the information stored or other stimuli involves attentional control (a central executive component) (Baddeley, 2012). Children with reading difficulties have been found to be disadvantaged both on measures of temporary storage only and on measures of combined storage and processing of linguistic content (Swanson, Zheng, & Jerman, 2009). As for specific comprehension difficulties, measures that tap processing in addition to storage are more predictive than measures of storage only (Carretti, Borella, Cornoldi, & de Beni, 2009). To complicate matters, inference making is also related to such measures of working memory (Cain, Oakhill, & Lemmon, 2004; Daneman & Merikle, 1996). However, work to date reports that variation in working memory does not fully explain the contribution of inference making to children's reading comprehension (Cain, Oakhill, & Bryant, 2004). Neither Cromley and Azevedo (2007), Ahmed et al. (2016) or Segers and Verhoeven (2016) included a measure of verbal working memory when assessing the effects of vocabulary and inference making on reading comprehension, but all authors acknowledged that working memory is a relevant covariate in the prediction of variation in reading comprehension.

Interestingly, the most powerful working memory measures for the prediction of differences in reading comprehension involve storage and processing of semantic information (Seigneuric, Ehrlich, Oakhill, & Yuill, 2000; Carretti, Borella, Cornoldi, & de Beni, 2009; Oakhill, Yuill, & Garnham, 2011; Nouwens, Groen, & Verhoeven, 2017). These measures could then in part predict reading comprehension because they place demands on participants' word knowledge and/or sentence

comprehension. This possibility underlines the importance of assessing the role of verbal working memory in a test format with minimal loadings on vocabulary and comprehension when investigating the contributions of linguistic components to reading comprehension.

The aim of this study was to investigate the relation between inference making and vocabulary knowledge in reading comprehension in young children, when a test of oral vocabulary and a comprehensive test of inference making is applied. Further, to take working memory into consideration as an explanatory variable. The research questions were:

1. Does inference making mediate the influence of vocabulary on reading comprehension?

2. Can verbal working memory account for the indirect effect of vocabulary on reading comprehension through inference making?

Method

Participants

Informed consent was obtained from 211 out of 308 students attending thirteen 6th grade classrooms (six schools). Students participated in group tests of reading comprehension and word reading correlates. From this sample, a smaller sample of 62 students (31 boys, 31 girls; mean age = 12 years and 7 months, *SD* = 4 months) was selected for individual assessments. Selection criteria for individual assessments included that students spoke Danish as their first language, students had no history of disabilities (such as dyslexia or ADHD), and students had word reading skills above -1.5 *SD* (individual assessments required several texts to be read aloud). Because an additional goal of the data collection was to establish small groups of poor comprehenders and good comprehenders (for analyses not reported here), students with these profiles were over-sampled. Although the sample of 62 students was not selected to be representative of all 6th graders, the average score of the selected sample corresponded to a z score of -0.1 for reading comprehension and a z score of 0.2 for word reading in the unselected sample. Thus, the average scores of the selected sample were very similar to the means of the unselected sample and therefore presumably comparable to the means of the population of 6th grade students in general.

Materials

Word reading. An estimate of word reading skills was derived to assess the independence of comprehension from word reading skills. The measure was a composite measure based on two group tests of phonological coding and orthographic choice (Nielsen & Petersen, 1993a). These tests were administered during the screening procedure. In the phonological coding test participants were required to find a pseudo-homophone among four non-words – that is to select the non-word that would *sound* like a real word if read aloud. This test has previously been shown to correlate strongly with non-word reading (r = .81) among dyslexic and non-dyslexic adults (Elbro, Nielsen, & Petersen, 1994) and with spelling of non-words (r = .48) (Nielsen & Petersen, 1993b). In the orthographic choice test participants were required to select the correct spelling of a word from four possible spellings. This test has previously been shown to correlate significantly with a standard test of sentence reading (r = .54) and spelling of irregular words (r = .62) (Nielsen & Petersen, 1993b). Both tests had a time limit and a maximum score of 38 (the number of correct items). The odd-even split-half reliabilities within the original unselected sample were high (.85 and .92).

Reading comprehension. An excerpt of the Diagnostic Reading Analysis test (DRA) (Crumpler & McCarty, 2004) was taken as the measure of reading comprehension. The excerpt comprised 4 narrative and 4 expository texts. According to the original manual of the DRA these texts were suitable for children between 10 and 14 years. Average text length was 70 words. The test included 46 open-ended questions, which were classified as literal (22), inferential/summative (14), predictive (6), and vocabulary (4) probes in the original test manual. Participants read texts out loud. The experimenter asked questions orally and participants answered orally with the text still available. During reading, the experimenter provided the correct pronunciation for specific words where a decoding difficulty arose. This procedure was undertaken to control for differences in word reading. Students' responses were recorded and subsequently scored in agreement with the test manual. Each correct answer was awarded 1 point. Internal consistency was acceptable (Cronbach's alpha = .73).

Inference making. Inference making abilities were assessed with an experimenter-developed test. Participants read short expository texts out loud and answered open-ended questions about the texts orally. Average text length was 51 words. The texts were designed using decoding vocabulary

appropriate for the age group (to minimise decoding failures). If a participant could not read aloud or mispronounced a word, the experimenter provided the correct pronunciation. A random sample of the read-alouds for the first 10 participants in the database showed a low error rate (< 1 %). The first part of the test consisted of 2 practice texts with subsequent questions. Then followed 24 texts and 32 questions (1-2 questions per text). Twenty-two questions were designed to elicit specific inferences. Ten 'filler' questions were literal questions. These were included to minimise the likelihood that participants would work out the specific character of the test (i.e., answers to questions are never stated explicitly in the text).

Students' responses were scored from audio recordings. For the inference questions, a correct answer was awarded 2 points. If a student provided an answer, which was judged to *potentially* lead to a correct answer the student was asked to elaborate on his or her answer (the experimenter asked: *Can you tell me more about that?*). These subsequent responses (after a prompt) received 1 point if correct, and no points if incorrect. The texts and questions displayed in the introduction section come from the experimental inference test. The literal questions were not analysed. However, it should be noted that the questions were very easy for students. All students were able to provide at least one correct piece of information to *every* literal question asked. This indicates low requirements for general comprehension skill of the test. The inference items were validated against the DRA using the mean percentage of correct answers for DRA texts with high inferential load (at least half of the questions for the texts were classified as either inferential/summative or predictive according to the test manual) and the mean percentage of correct answers for DRA texts with low inferential load (the remaining texts). The correlation between the inference test score and comprehension of texts expected to have a high inferential load ($r = .60$, $p < .01$) was stronger than the correlation with comprehension of texts expected to have low inferential load ($r = .37$, $p < .01$). The difference was significant ($z = 2.34$, $p < .05$).

Four inference items were found to be unreliable. They allowed a sensible answer that conflicted with the target inference but which was in accordance with general background knowledge. These

four items were excluded in the analyses. Cronbach's alpha was acceptable (.74) for the remaining items.

Vocabulary. A Danish adaptation of the Peabody Picture Vocabulary Test (PPVT) (Dunn, 1959; Nielsen, 2008) provided a measure of vocabulary knowledge. For each item, the participant was asked to select one out of four pictures that best illustrated the meaning of a word spoken aloud by the experimenter. When a participant had failed to identify the correct picture for six items (within the range of 8 items) the test was terminated. Cronbach's alpha was calculated for the majority of items (67 %), which were attempted by a majority of participants (64 %). This was acceptable (= .72).

Working memory. The Backward Digit Span subtest from the Clinical Evaluation of Language Fundamentals (CELF-4) (Semel, Wiig, & Secord, 2013) was administered as a measure of verbal working memory. In contrast to forward digit span tasks, backward digit span tasks require not only storage of information (short term memory), but also processing of that information (central executive) (Reynolds, 1997; Swanson, Howard, & Sáez, 2006; Swanson, Zheng, and Jerman, 2009). The experimenter spoke out strings of digits and the participant recalled these strings in the reverse order, thereby simultaneously storing and processing the digits. The number of digits was gradually increased and the test was terminated when the participant failed to recall two consecutive strings of a certain length. One point was awarded for each correctly recalled string of numbers. A reliability coefficient of .71 for the relevant age group (12;00-12;11) is reported in the test manual.

Procedure

The participants were assessed individually over two sessions. Assessments took place after school either at the school of the participant or at the University of Copenhagen. Participants received a gift card for cinema tickets at the end of the second session.

Results

Complete data sets were obtained and analysed from 53 participants. All score distributions had values of skewness and kurtosis within +/- 0.79. The maximum z values for skewness and kurtosis were 2.15 and 1.21, which are considered acceptable values for a normal distribution with small to medium-sized samples (Kim, 2013). Descriptive statistics and correlations are reported in Table 1.

[TABLE 1]

Reading comprehension and the word reading composite did not correlate significantly. This result confirmed that the assessment procedure of the reading comprehension test had efficiently controlled for word reading differences. Inference making, vocabulary and verbal working memory were all significantly correlated with reading comprehension. As expected, both vocabulary and verbal working memory were also significantly correlated with inference making. These correlations were all moderate to large. The correlation between vocabulary and verbal working memory was low and did not reach significance (r = .25, p = 0.07). This indicates that the verbal working memory test did not tap semantic word knowledge to any significant degree.

As noted above, there was a significant correlation between vocabulary and reading comprehension. To determine if inference making mediated the influence of vocabulary on reading comprehension, a mediation analysis was conducted using the extension programme *Process* for SPSS (Hayes, 2013). Standardised scores were used in the analyses. A bias-corrected bootstrap interval for the indirect effect based on 5000 bootstrap samples was constructed. The indirect effect of an independent variable through a mediator is considered statistically significant if zero is not included in the bias-corrected confidence interval (Hayes, 2013). The results of this simple mediation analysis with vocabulary as independent variable and inference making as mediator, are shown in Figure 2. The indirect effect of vocabulary on reading comprehension through inference making (*ab*) was significant. The direct effect of vocabulary on comprehension (*c'*), which controlled for the effect of inference making, was not significant. Inference making therefore fully mediated the effect of vocabulary on reading comprehension.

[FIGURE 2]

To assess whether the data supported the alternative mediation hypothesis we reran the analysis, this time treating inference making as the independent variable and vocabulary as the mediator. This analysis provided only weak support for the alternative hypothesis. The indirect effect of inference making through vocabulary (*ab* = .10) was not considered statistically significant since zero was in the biased-corrected confidence interval (BC CI [-0.06, 0.28]).

To answer the second research question regarding the role of working memory, a mediation analysis was conducted with verbal working memory as a covariate. Results are shown in Figure 3. The analysis showed that, even after controlling for verbal working memory, the indirect effect of vocabulary on reading comprehension through inference making was significant. The effect size was only slightly reduced (from 0.25 to 0.22). Vocabulary remained a significant predictor of inference making when verbal working memory was entered as a covariate, but verbal working memory did not contribute significantly to inference making when controlling for the effect of vocabulary. However, verbal working memory continued to contribute to reading comprehension over and above inference making and vocabulary.

[FIGURE 3]

Discussion

The results of this study converge with those reported by Cromley and Azevedo (2007), Ahmed et al. (2016) and Segers and Verhoeven (2016) in that they indicate that the effect of vocabulary on reading comprehension is *at least partly* mediated by inference making. This study, therefore, provides a replication of those earlier findings with a different sample and different, more comprehensive, assessments of the critical independent variables. In addition, the current study extends those earlier findings by taking into account the possible alternative mediation hypothesis and by elucidating the role of working memory on reading comprehension, in relation to inference making and vocabulary. Together the results of these studies constitute a first step in the process of confirming the mediating role of inference making in reading comprehension. Being correlational studies, a next obvious step in the process is the conduction of a longitudinal study.

In contrast to the previous studies we did not find a direct effect of vocabulary on reading comprehension. This could reflect limited power due to our small sample size (Hayes, 2013). It might also have to do with the sensitivity of the specific vocabulary measure, which can be regarded as a measure of vocabulary breadth (number of words known). Research suggest that vocabulary measures that capture the number of semantic features known as well as word connections (often defined as measures of ‘vocabulary depth’) are more strongly related to reading comprehension and inference

making than traditional tests of vocabulary breadth (Cain & Oakhill, 2014; Ouellette, 2006). However, the vocabulary measure in the study of Cromley and Azevedo (2007) might also be classified as a measure of vocabulary breadth, and so might some of the tasks utilized by Ahmed at al. (2016). Vocabulary breadth and depth lie on a continuum and as a result, tasks that have been classified as measures of depth by some authors have been classified as breadth measures by others (compare Tannenbaum, Torgesen, & Wagner (2006) with Ouellette (2006)). An important aim for future studies appears to include valid measures of both breadth and depth of vocabulary and shed light on the interplay between these different aspects of vocabulary and inference making in reading comprehension.

A possible limitation of the study lies in that the sample was not representative of all 6th graders. Even though the sample means were probably not different from the population means, the sample variation was probably larger than in the population because poor and good comprehenders were over-sampled. This may have increased the study's sensitivity to associations between reading comprehension and related abilities, i.e., vocabulary, inference making, and verbal working memory. However, it is unlikely that increased variability influenced the *relations* between the weights of the related abilities.

Although inference making is considered to be dependent on working memory, entering working memory as a covariate in the mediation analysis did not significantly alter the indirect effect of vocabulary through inference making. This result is in line with previous research (Cain, Oakhill, & Bryant, 2004; Currie & Cain, 2015) and shows that inference making is not determined solely by working memory resources. The applied working memory measure was one with minimal semantic loadings. One might expect that other working memory measures are able to account for more shared variance with inference making. However, as noted, working memory measures that put demands on students' word knowledge and/or sentence comprehension may obscure the interpretation of the contributions of linguistic components to reading comprehension.

Two important limitations with respect to the influence of working memory should also be noted. First of all, the texts in the inference test were short. The different pieces of information that were

critical to the targeted inferences were often found in two or more *adjacent* sentences. This may have reduced the impact of working memory differences on inference making. Second, participants had access to the texts in the inference test while they answered the questions. This may also have reduced the influence of working memory. However, it should be noted that access to texts while answering questions also applied to the reading comprehension task and, on this task, working memory was found to have a significant independent effect.

The test of inference making had some similarities to the test of reading comprehension. In both tests participants were asked to read texts and answer open-ended questions about the texts. A reasonable question to ask is therefore to what extent the inference test assessed a construct different from general reading comprehension. It is inevitable that an inference test (targeting the kind of inference making that is important in reading comprehension, e.g., inferences that establish referential and causal coherence) will entail some aspects of general comprehension. However, the experimental inference test employed in this study was, contrary to the publisher-developed test of reading comprehension, developed for the sole purpose of assessing inference making. It goes without saying that the score on the inference test was calculated exclusively on the basis of the questions designed to tap specific inferences. A much wider variety of question types fed into the reading comprehension score – in accordance with general reading comprehension being a multi-faceted skill. Further, students' apt responses to the literal filler questions in the inference test indicated that these texts did not place notable demands on literal comprehension (a facet of general comprehension that was included in the reading comprehension score). The validity of the inference test was further supported by the fact that inference scores were significantly more strongly related to comprehension of texts with high inferential demands (determined on the classification of items in the test manual) than texts with low inferential demands.

Importantly, the two tests in question were differentially associated with the verbal working memory measure, which also indicates that the tests assess different constructs. In accordance with the bulk of research literature on reading comprehension and verbal working memory (e.g., Daneman & Carpenter, 1980) reading comprehension was significantly associated with working memory, even

when controlling for inference making and vocabulary. In contrast, the association between working memory and inference making was no longer significant once vocabulary was controlled. This result concurs with recent research showing that vocabulary mediates the effect of working memory on children's inference making (Currie & Cain, 2015).

To further work around the possible overlap between inference making measures and measures of reading comprehension, future studies might follow Segers and Verhoeven's (2016) example and employ a proxy of inference making such as a measure of reasoning detached from text comprehension. However, such studies will have to investigate the amount of variance in reading comprehension that is shared by conventional measures of inference making and measures of logical reasoning to confirm a common construct.

As already pointed out the present study was correlational by nature and does therefore not imply causal relationships between vocabulary, inference, and reading comprehension. Longitudinal studies and training studies will be needed to resolve causal issues. Results from the present study are however in agreement with other recent work, suggesting that gains in vocabulary in itself are not sufficient to improve general reading comprehension (Elleman et al., 2009; Apthorp, et al., 2012). As mentioned, Beck et al. (1982) and Vadasy et al. (2015) were able to show a small, but significant, effect of vocabulary instruction on standardized reading comprehension measures. The vocabulary instruction in these two studies was Rich Instruction, which includes a strong focus on interrelationships between words and reasoning practice activities. Apthorp et al. (2012) also employed instruction with these elements, but did not find a transfer effect to general reading comprehension. Results from the present study suggest, that it might be possible to improve the outcome of such vocabulary instruction on reading comprehension if the interplay between vocabulary and inference making is accentuated during instruction.

To conclude, inference making was found to mediate the effect of vocabulary on comprehension. This indicates that word meanings are important for comprehension in the context of inference making. Two reasons for this have been highlighted: Inference making taps semantic relations among

words, and precise word meanings in texts are selected on the basis of context. The indirect effect of vocabulary through inference making was independent from verbal working memory.

References

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology*, *44–45*, 68–82. doi: 10.1016/j.cedpsych.2016.02.002

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The Contribution of Attentional Control and Working Memory to Reading Comprehension and Decoding. *Scientific Studies of Reading*, *18*, 325–346. doi: 10.1080/10888438.2014.902461

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. doi: 10.1126/science.1736359

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*, 1–29. doi: 10.1146/annurev-psych-120710-100422

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*, 506–521. doi: 10.1037//0022-0663.74.4.506

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal*, *11,* 489–503. doi: 10.1023/A:1008084120205

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114,* 647–662. doi: 10.4074/S0003503314004035

Cain, K., Oakhill, J., & Bryant, Peter. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills, *Journal of Educational Psychology*, *96,* 31–42. doi: 10.1037/0022-0663.96.1.31

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*, 671–681. doi: 10.1037/0022-0663.96.4.671

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*, 246–251. doi: 10.1016/j.lindif.2008.10.002

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786. doi: 10.3758/BF03196772

Cromley, J. G. & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*, 311–325. doi: 10.1037/0022-0663.99.2.311

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

Cromley, J. G., & Wills, T. W. (2016). Flexible strategy use by students who learn much versus little from text: transitions within think-aloud protocols. *Journal of Research in Reading*, *39*, 50–71. doi: 10.1111/1467-9817.12026

Crumpler, M., & McCarty, C. (2004). *Diagnostic reading analysis: Manual*. London, UK: Hodder & Stoughton.

Cunningham, A. E., & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction*, *1*, 137–149. http://www.nifdi.org/research/journal-of-di

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, *137*, 57–75. doi: 10.1016/j.jecp.2015.03.005

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

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 422–433. doi: 10.3758/BF03214546

Dunn, L. (1959). *The Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading*, *17*, 435–452. doi: 10.1080/10888438.2013.774005

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, *44*, 203–226. doi: 10.1007/BF02648162

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*, 1–44. doi: 10.1080/19345740802539200

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

Graesser, A., Wiemer-Hastings, P., & Wiemer-Hastings, K. (2001). Constructing inferences and relations during text comprehension. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 249–226). Amsterdam: John Benjamins Publishing Company.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*, 52–54. doi: 10.5395/rde.2013.38.1.52

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

Kintsch, W. & Rawson, K. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209–226). Malden, MA: Blackwell.

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language*, *33*, 175–202. doi: 10.1006/jmla.1994.1009

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*, 237–270. doi: 10.2307/1162893

Nation, K. (2009). Reading comprehension and vocabulary: What's the connection? In R. K. Wagner, K., C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 176–194). New York, NY: Guilford Press.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research: JSLHR*, *47*, 199–211. doi: 10.1044/1092-4388(2004/017)

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*, 342–356. doi: 10.1111/j.1467-9817.2004.00238.x

Nielsen, I., & Petersen, D., K. (1993a). *Diavok, en diagnostisk læse- og stavetest* [Diagnostic reading and spelling test]. Copenhagen: AOF's Landsforbund.

Nielsen, I. & Petersen, D. K. (1993b). *Udvikling og afprøvning af et diagnostisk testmateriale til voksne dyslektikere* [Development and testing of a diagnostic test material for dyslexic adults] [unpublished master's thesis No: 71000305]. University of Copenhagen, Copenhagen.

Nielsen, J. B. (2008). *Det mentale leksikon og testning af receptivt ordforraad, aendring af Peabody-testen* [The mental lexicon and assessment of receptive vocabulary, an adaptation of the Peabody test] [unpublished master's thesis No: 4764-088-4]. University of Copenhagen, Copenhagen.

Nouwens, S., Groen, M. A., & Verhoeven, L. (2017). How working memory relates to children's reading comprehension: the importance of domain-specificity in storage and processing. *Reading and Writing: An Interdisciplinary Journal*, *30*, 105–120. doi: 10.1007/s11145-016-9665-5

Oakhill, J., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*, 91–121. doi: 10.1080/10888438.2010.529219

Oakhill, J., Yuill, N., & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension. *International Electronic Journal of Elementary Education*, *4*, 83–106. http://www.iejee.com/

Osana, H. P., Lacroix, G. L., Tucker, B. J., Idan, E., & Jabbour, G. W. (2007). The impact of print exposure quality and inference construction on syllogistic reasoning. *Journal of Educational Psychology*, *99*, 888–902. doi: 10.1037/0022-0663.99.4.888

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

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell.

Perfetti, C., Yang, C.-L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*, 303–318. doi: 10.1002/acp.1419

Reynolds, C. R. (1997). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology*, *12*, 29–40. doi: 10.1093/arclin/12.1.29

Segers, E., & Verhoeven, L. (2016). How logical reasoning mediates the relation between lexical quality and reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *29*, 577–590. doi: 10.1007/s11145-015-9613-9

Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *13*, 81–103. doi: 10.1023/A%3A1008088230941%I

Semel, E., Wiig, E. H., & Secord, W. (2013). *CELF-4 Clinical evaluation of language fundamentals (4th ed). Dansk vejledning* [Danish manual] Enschede, NLD: Pearson.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities*, *39*, 252–269. doi: 10.1177/00222194060390030501

Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities; Austin*, *42*, 260–287. doi: 10.1177/0022219409331958

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, *10*, 381–398. doi: 10.1207/s1532799xssr1004_3

Vadasy, P. F., Sanders, E. A., & Logan Herrera, B. (2015). Efficacy of rich vocabulary instruction in fourth- and fifth-grade classrooms. *Journal of Research on Educational Effectiveness*, *8*, 325–365. doi: 10.1080/19345747.2014.933495

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Hillsdale, NJ: Lawrence Erlbaum Associates.

van den Broek, P., White, M., Kendeou, P., & Carlson, S. (2009). Reading between the lines: Developmental and individual differences in cognitive processes in reading comprehension. In R. K. Wagner K., C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 107–123). New York, NY: Guilford Press.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*, 8–25. doi: 10.1080/10888438.2011.536125

Table 1.

Descriptive statistics and correlation coefficients between the variables

Measure	Correlations				*M*	*SD*	Min.	Max.
Reading comprehension (max. = 46)					36.9	4.6	25	46
Inference making (max. = 36)	.58**				25.9	6.0	11	36
Vocabulary (max. = 121)	.44**	.52**			74.2	12.6	57	104
Verbal working memory (max. = 14)	.41**	.32*	.25		5.0	1.4	2	8
Word reading composite (z-score)	.14	.24	.18	.50**	0.00	0.76	-1.47	1.66

Note. *$p < .05$; **$p < .01$.

[…] The king had some banks of earth made around Copenhagen and he built fortresses along the borders of the country. Here people stayed ready to fight. During the same period the king also built a new naval port in Copenhagen. In the facility weapons and food were being loaded. […] Question: What went on in the new naval port?	[…] In the course of time they learned how to bake bread. The first loaves were flat loaves, which reminded of crispbread. At one point in antiquity women in Egypt began to use water from The Nile for bread. The water from The Nile contains yeast. Question: What happened to the loaves when women in Egypt began to use water from The Nile for the bread?

Figure 1. Examples from inference task. The examples are translated from Danish.

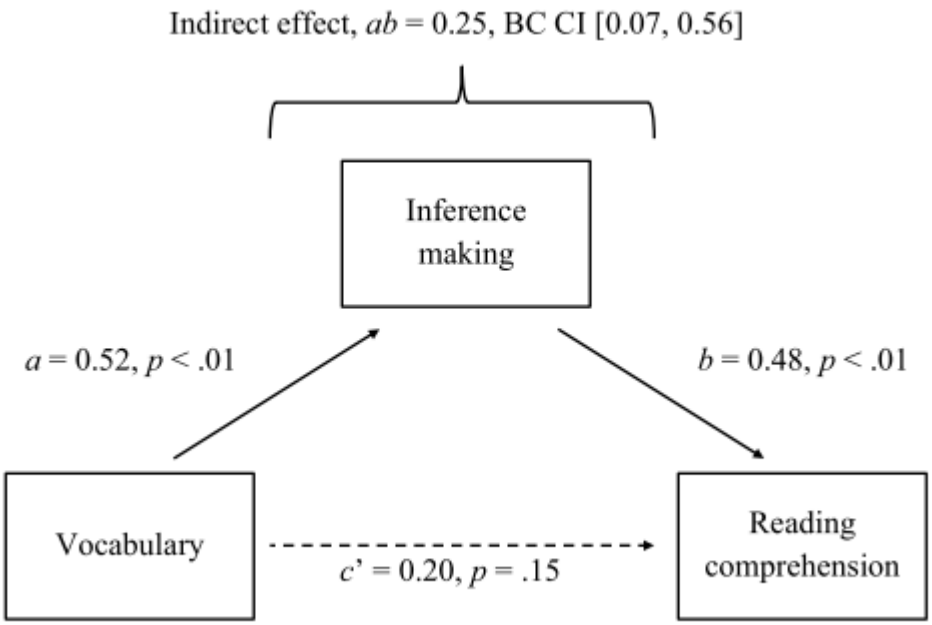

Figure 2. Results of the simple mediation analysis.

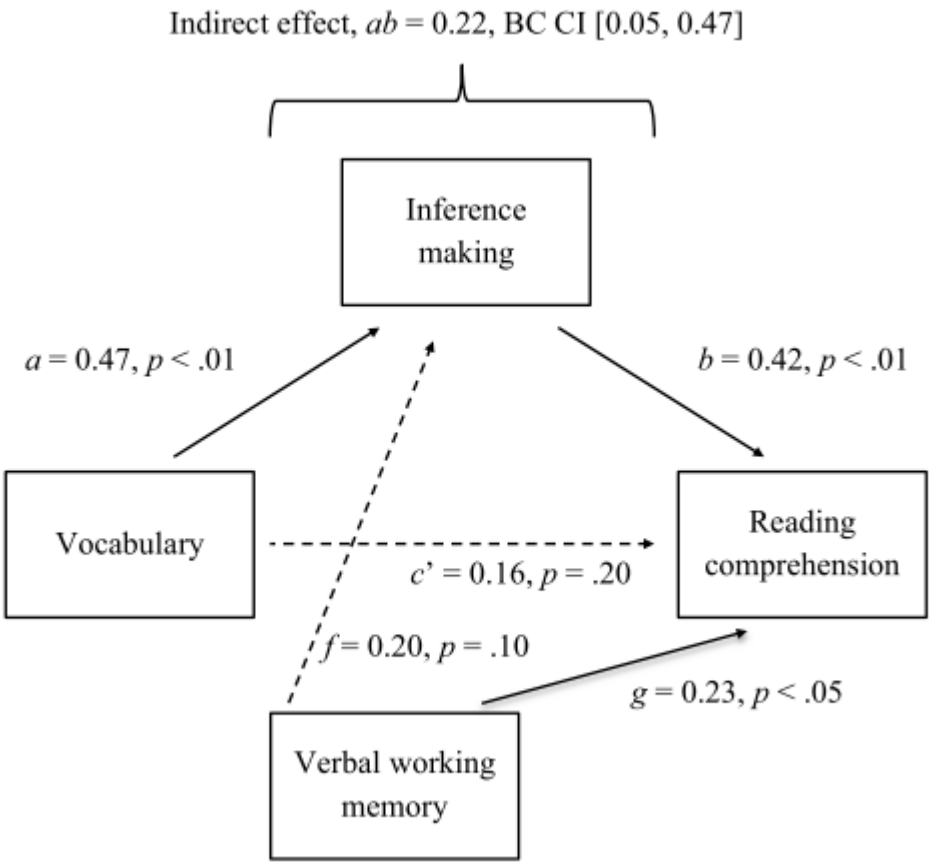

Figure 3. Results of the mediation analysis with verbal working memory as a covariate.