Journal of Experimental Child Psychology 137 (2015) 57–75

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children's inference generation: The role of vocabulary and working memory

Nicola Kate Currie, Kate Cain *

Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK

ARTICLE INFO

Article history:
Received 21 June 2014
Revised 5 March 2015

Keywords:
Inference
Local coherence
Global coherence
Working memory
Vocabulary
School-aged children

ABSTRACT

Inferences are crucial to successful discourse comprehension. We assessed the contributions of vocabulary and working memory to inference making in children aged 5 and 6 years ($n = 44$), 7 and 8 years ($n = 43$), and 9 and 10 years ($n = 43$). Children listened to short narratives and answered questions to assess local and global coherence inferences after each one. Analysis of variance (ANOVA) confirmed developmental improvements on both types of inference. Although standardized measures of both vocabulary and working memory were correlated with inference making, multiple regression analyses determined that vocabulary was the key predictor. For local coherence inferences, only vocabulary predicted unique variance for the 6- and 8-year-olds; in contrast, none of the variables predicted performance for the 10-year-olds. For global coherence inferences, vocabulary was the only unique predictor for each age group. Mediation analysis confirmed that although working memory was associated with the ability to generate local and global coherence inferences in 6- to 10-year-olds, the effect was mediated by vocabulary. We conclude that vocabulary knowledge supports inference making in two ways: through knowledge of word meanings required to generate inferences and through its contribution to memory processes.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

* Corresponding author.
E-mail address: k.cain@lancaster.ac.uk (K. Cain).

http://dx.doi.org/10.1016/j.jecp.2015.03.005
0022-0965/© 2015 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Skilled comprehenders make sense of written and spoken language by constructing a coherent memory-based representation of the state of affairs described by the text, commonly referred to as a situation model (Kintsch, 1988). The text does not always explicitly state all of the information needed for coherence. Therefore, readers and listeners regularly make inferences to integrate information within the text and to fill in details that are only implicit. These inferences are incorporated into the situation model that the comprehender constructs and result in a more accurate and complete understanding of the text (Graesser, Singer, & Trabasso, 1994). Although children make inferences from an early age, they do not typically make as many as do older children and adults (e.g., Ackerman, 1986; Casteel, 1993). Our focus in this study was to understand better why this is the case by examining the role played by two critical factors related to young children's inference generation: vocabulary and working memory. Furthermore, we explored whether these factors make different contributions to distinct types of inference. A greater understanding of the factors that support early inference making will inform models of comprehension development and also the literacy curriculum and targeted intervention programs for children with weak inference-making and comprehension skills.

When considering different types of inference, one of the key distinctions that can be made is between inferences that establish local coherence and those that establish global coherence (Graesser et al., 1994). Inferences that are necessary for local coherence typically involve the integration of separate propositions within the text and are usually cued by a pronoun, synonym, or category exemplar; for example, "He finished the orange juice quickly. The drink was very refreshing." Local coherence inferences often require a mapping between related words, for example, between synonyms or category exemplar pairings as in the example above. In contrast, inferences necessary for global coherence can involve inferring goals that motivate particular actions or establishing an overall theme of a text, and this often relies on the ability to connect ideas that are not explicitly signaled by a single word and that can be distributed throughout the text. For example, readers can infer the likely setting of a story through links between semantically related concepts such as "building sandcastles," "paddling in the water," and the presence of a "pier" (which together indicate that the setting is the seaside). In this way, inferences that are necessary for global coherence typically draw more heavily on information that is external to the text than do local coherence inferences (Cain & Oakhill, 1999, in press). We note that it is likely that local and global coherence inferences are not truly categorical distinctions and that, rather, different inferences draw on information in the text and background knowledge to differing degrees (Florit, Roch, & Levorato, 2011). Furthermore, some authors distinguish local and global coherence inferences in terms of distance between elements of the text (e.g., McKoon & Ratcliff, 1992), but we did not manipulate this feature of our texts. Critically, both local and global coherence inferences are necessary for a full understanding of the text, in line with Cain and Oakhill's (1999) study of children and Long and Chong's (2001) study of adults.

Adults routinely make inferences to establish local and global coherence (e.g., Albrecht & O'Brien, 1993; Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Nicol & Swinney, 1989; Sanders & Noordman, 2000). With regard to local coherence inferences, children are capable of making them from an early age, but developmental improvements are clear. Ackerman (1986) found that 6-year-olds were as sensitive as 9-year-olds to the need to establish local coherence in short texts but made fewer local coherence inferences than did 9-year-olds when required to support comprehension. Barnes, Dennis, and Haefele-Kalvaitis (1996) demonstrated substantial developmental improvements in the ability to make local coherence inferences between 6 and 11 years of age, with smaller gains thereafter up to 15 years (the oldest age group in their study). Likewise, research examining children's ability to generate inferences to establish global coherence demonstrates early sensitivity as well as developmental gains, with 4-year-olds making fewer inferences about narrative themes and characters' goals than 6-year-olds (Lynch et al., 2008).

These studies provide clear evidence that, despite early sensitivity to the need to establish both local and global coherence when processing text, significant gains in inference-making ability occur during early to middle childhood. When we examine the wider literature on inference making in adults, as well as in children, we find evidence that vocabulary and working memory influence an individual's inference-making ability. We next turn our attention to a consideration of these skills and how they might be related to developmental improvements in inference making.

Obviously, text comprehension could not occur without knowledge of individual word meanings, and for that reason vocabulary is routinely shown to be related to general measures of reading (Oakhill & Cain, 2012) and listening comprehension (Florit, Roch, Altoe, & Levorato, 2009). Vocabulary is also significantly correlated with measures of inference making in both children (Lynch et al., 2008; Oakhill & Cain, 2012) and adults (Dixon, LeFevre, & Twilley, 1988). For local coherence inferences, knowledge of word meanings may be particularly important when a synonym, paraphrase, or category member refers back to an earlier mentioned object (Perfetti, Yang, & Schmalhofer, 2008). In addition to knowledge of specific word meanings, the production of some inferences also draws heavily on background knowledge and the interrelations between words (Cain & Oakhill, 2014; Casteel, 1993). In the case of global coherence inferences, one cannot infer that a furry animal that barks and likes going for walks is a dog unless one possesses the requisite knowledge about dogs and their characteristics. Thus, not just knowledge of individual word meanings but also rich semantic networks with robust connections between the meanings of words associated by topic may be important for easy and accurate inference making. Therefore, there are clear mechanisms through which vocabulary may support both local and global coherence inference making.

Working memory refers to the memory systems used for the simultaneous storage and processing of information (Baddeley & Hitch, 1974). Measures of verbal working memory (sometimes referred to as verbal complex memory span) are more strongly related to reading comprehension in young children than are measures that simply tap storage of verbal information (or short-term span) (Leather & Henry, 1994). Furthermore, working memory explains unique variance in general measures of listening and reading comprehension after controlling for vocabulary in children between 4 and 10 years of age (Cain, Oakhill, & Bryant, 2004; Florit et al., 2009; Seigneuric & Ehrlich, 2005).

Working memory may be particularly important for inference generation because a reader or listener needs to maintain activation of previously processed information while relating this to the piece of text currently being processed. Younger children might routinely fail to do so if their working memory capacity limits the amount of information that they can store when processing text. This may account for the observation that younger children tend to process text in a piecemeal manner, not always making connections between ideas, particularly if they are not presented in succession (Schmidt & Paris, 1983). However, studies that have included assessments of working memory and either inference making or more general reading comprehension within a developmental framework report contradictory results. Chrysochoou and Bablekou (2010) found a reduction in the influence of working memory on inference making between 5 and 9 years of age, whereas other work suggests that the relation between working memory and reading comprehension either stays constant between 7 and 10 years of age (Seigneuric & Ehrlich, 2005) or increases during that period (Cain et al., 2004). One factor influencing the strength of any relation between inference and working memory may be the nature of the materials; the influence of working memory on inference making for 9- and 10-year-olds is strongest when the pieces of information supporting the inference need to be integrated over large units of text (Cain, Oakhill, & Lemmon, 2004).

In the current study, we were interested in the contributions made by both vocabulary and working memory to local and global coherence inferences in different age groups. Recent work supports this focus by showing that from as young as 4 to 6 years, higher level comprehension skills such as text integration (particularly relevant for local coherence inferences) and knowledge accessibility (important for both inference types but particularly for global coherence inferences) emerge as separate skills (Hannon & Frias, 2012). To date, there is only one published study that has explored the relative influences of both working memory and vocabulary to different types of inference making (Chrysochoou, Bablekou, & Tsigilis, 2011). The two inference types studied by Chrysochoou and colleagues (2011) were required to establish coherence in the text. They broadly map onto the local and global coherence distinction that we focused on here. We note that the authors referred to the latter as elaborative inferences, but on examination these appear to be necessary to ensure a full and coherent understanding of the text as intended by the authors of the original article (Cain & Oakhill, 1999). Chrysochoou and colleagues (2011) found small, but significant, associations between complex working memory span and inferences that were required for local coherence in their sample of 9-year-olds. However, they found a much stronger association between working memory and inferences required to establish global coherence. The relations between their measures of short-term span (e.g., measures of the phonological loop) and both types of inference were weak.

Critically, Chrysochoou and colleagues (2011) demonstrated that vocabulary fully mediated the relations between working memory and the inferences required for local coherence but only partially mediated the relations between working memory and the inferences required to establish global coherence. One possible reason for this distinction is the way in which vocabulary can support memory. The theory is that children and adults with richer vocabulary knowledge have more accurate and available representations of words in their long-term memory than those with poorer vocabulary knowledge, and this better knowledge supports accurate maintenance of information in verbal working memory (Nation, Adams, Bowyer-Crane, & Snowling, 1999; Walker & Hulme, 1999). Thus, vocabulary may be important for inference making in two ways. First, vocabulary knowledge is important because inferences involve word knowledge; local coherence inferences involve mapping between synonyms and category exemplars, and global coherence inferences tap knowledge about the interrelations between word meanings. Second, vocabulary knowledge may support inference making because it can provide a boost to accurately maintain the contents of working memory, necessary to aid the integration of information from different parts of the text. In relation to the findings of Chrysochoou and colleagues (2011), the full mediation of working memory by vocabulary for local coherence inferences may reflect the lower memory demands associated with this type of inference, which involve integration between successive sentences, and the partial mediation for global coherence inferences may reflect the higher memory demands of global coherence inferences, which involve integration of ideas throughout the text and also with background knowledge external to the text.

Another possibility for this distinction is that Chrysochoou and colleagues (2011) used a single measure of receptive vocabulary, which is regarded by some as a measure of the number of words known and which does not tap deeper knowledge of the interrelations between words (Tannenbaum, Torgesen, & Wagner, 2006). We assessed vocabulary in two ways to achieve a more complete assessment of this complex construct; we used the same single word comprehension measure as Chrysochoou and colleagues and also a measure that assessed knowledge of word networks (semantic fluency). We believed that this broader conceptualization of vocabulary knowledge would provide a robust test of its relation to inference making.

The current study sought to build on previous research on children's inference making by looking specifically at the distinction between inferences required to establish local and global coherence (a distinction that has not been examined in previous work with younger age groups of children) with the aim to determine how vocabulary and working memory influence this ability in 6- to 10-year-olds. Our review of the literature indicates that both vocabulary and working memory will be associated with inference making in this age range. We predicted that vocabulary would be more strongly associated with global coherence inferences than with local coherence inferences because the former rely on richer and better connected semantic networks, in line with Cain and Oakhill (2014). We also predicted that memory would be more strongly associated with global coherence inferences than with local coherence inferences because global coherence inferences involve maintaining activation of a larger amount of text.

Regardless of their age, all children in our sample were presented with the same texts, which were read aloud to them by the assessor. It is widely believed that the same skills underpin comprehension of written and spoken text (Hoover & Gough, 1990). This is supported by research demonstrating the role of inference for each (Barnes et al., 1996) and strong relations between performance on reading and listening comprehension tasks when word reading ability is suitably controlled (Cain, Oakhill, & Bryant, 2000; Kendeou, van den Broek, White, & Lynch, 2009; Stothard & Hulme, 1992). Because we used the same texts, we expected to find developmental improvements in inference making and the strongest relation between memory capacity and inference making in the youngest age group, where the processing demands would be greatest. In general, we expected stronger associations between inference making and measures of complex memory span (tasks that tap both storage and processing) than between inference making and short-term memory span measures (which tap only storage), similar to the results found for 8- and 9-year-olds by Chrysochoou and colleagues (2011). There are contradictory findings for the roles of vocabulary and memory; studies of 8- to 11-year-olds find a unique role for working memory in the prediction of concurrent reading comprehension that is independent of vocabulary (Cain et al., 2004; Seigneuric & Ehrlich, 2005), whereas other work suggests that working memory has a unique role in the prediction of reading comprehension, over and above vocabulary knowledge, for 5- and 7-year-olds but not for 9-year-olds (Chrysochoou & Bablekou, 2010; Florit et al., 2009).

A key aim of this study was to determine whether vocabulary and working memory predicted unique variance in inference making or whether any relation between working memory and inference making was mediated by vocabulary as found by Chrysochoou and colleagues (2011; see also Dixon et al., 1988, for work with adults). This finding is in contrast to studies of discourse comprehension that identify a unique contribution for working memory on reading and listening comprehension independent of vocabulary (Cain et al., 2004; Florit et al., 2009; Seigneuric & Ehrlich, 2005). Full mediation would indicate that the relation between working memory and inference making is due to the support that vocabulary provides for maintaining activation of relevant information in memory. Thus, our analyses were constructed to determine whether this relation exists across all ages for both types of inference.

Method

Participants

The participants were 130 children from schools in the northwest of England. Of this total sample, 44 children were enrolled in Year 1 classrooms (5–6 years of age, *M* = 6;2 [years;months], *SD* = 4 months, range = 5;7–6;8; 26 boys and 18 girls), 43 were from Year 3 classrooms (7–8 years of age, *M* = 8;3, *SD* = 3 months, range = 7;9–8;8; 21 boys and 22 girls), and 43 were from Year 5 classrooms (9–10 years of age, *M* = 10;2, *SD* = 4 months, range = 9;8–10;8; 24 boys and 19 girls). Consent was obtained from headteachers and parents, and assent was received from children, prior to each assessment session. The schools served socially mixed catchment areas. All children spoke British English as their first language, and children with a statement of special educational needs did not take part in the study.

Materials and procedure

All children completed assessments of inference making, vocabulary, and memory. They were assessed individually over four sessions, with each session lasting no longer than 15 min.

Local and global coherence inferences

Each child listened to four short stories modified from materials developed for another project (Language and Reading Research Consortium, 2015) that were read aloud by the experimenter. The stories were 148 to 161 words in length. Each story had an episodic structure that began with a setting (to introduce characters, time, and/or place) and had categories of story units, including events, goals, attempts, outcomes, and/or reactions (Stein & Glenn, 1982). The Coh-Metrix Text Easability Assessor (Graesser, McNamara, & Kulikowich, 2011) metrics showed that the texts had high narrativity (the extent to which the text is story-like, *M* = 63%), had a high number of concrete and imageable content words (*M* = 94%), and were syntactically simple (*M* = 94%) but had low referential cohesion (*M* = 35%). The latter was to be expected because the texts were constructed to require inferential processing for adequate comprehension. Children were instructed to listen carefully to the stories so that they could answer the memory questions after each one. At the end of each story, children were asked four questions that tapped the ability to make local coherence inferences and four questions that tapped global coherence inferences. For each story, the four local coherence inferences required the listener to integrate information from two sentences in the text and the four global coherence inferences required the listener to understand motivations and infer themes, settings, or central character identity.

Table 1
Story extract and corresponding inference questions.

Billy, Susie, and their Mum had gone out for the day. Billy spent the morning building a **sandcastle** near the **water**. Mum sat on their large **beach towel** and read a book. Susie wanted to go for a **swim**. She put her feet in the **sea**, but the water felt too cold. Susie went and sat down next to Mum instead. Mum had packed a big bag full of books and games. Susie found her story book and started to read.
Local coherence inference question: Where did Susie find her book?
Answer: In the bag
Prompt if incomplete answer: If partial response such as "Mum packed them," prompt with "But where was the book?"
Global coherence inference question: Where were Billy and his family?
Answer: Beach
Prompt if incomplete answer: If child responds "Out for the day," prompt with "Can you tell me where they were?"
Background Knowledge Check 1: "If there was the sea and sand, where could you be?"
Answer: Seaside, beach, coast
Background Knowledge Check 2: "Is there sea and sand at the beach?"
Answer: Yes

Note. Underlined text indicates information used for local coherence inference, and bold text indicates information that could be used for global coherence inference.

Examples are provided in Table 1. A second rater categorized each of the 32 questions. There was 94% agreement for the local versus global coherence distinction. The two questions categorized differently were resolved through discussion. Cronbach's alpha on this task for this sample of children was adequate ($\alpha = .77$).

The order of the questions followed the order of information presentation in the story.[1] If a child gave an incomplete or vague answer, the experimenter prompted the child. This was either a repetition of the question or encouragement to be more specific (see Table 1 for examples).

Each inference question was scored as correct, correct after a prompt, or incorrect. Correct responses were awarded 2 points for a full response and 1 point for a partial response. When a child did not provide an acceptable response to a global coherence question, the child's background knowledge for that information was checked with a follow-up question (see Table 1). Although both local and global inferences draw on relevant background knowledge, the global coherence inferences examined in the current study were more knowledge dependent in that they drew on a range of facts to work out a setting or a character's identity. Therefore, we made additional checks that children possessed the necessary background knowledge for this inference type so that we could determine whether knowledge differences were the source of any developmental differences found. There were very few instances where children did not have the required background knowledge; only one child in each year group answered one background knowledge question incorrectly (<0.3%). An adjusted total score was calculated taking into account only those items for which the child possessed the relevant background knowledge. This was a proportionate score based on the child's actual score and the maximum possible score (for only those items where the background knowledge was known). To enable ease of comparison, the raw correct scores for the local coherence questions are reported as proportions correct.

Vocabulary

Each child completed two measures of vocabulary, one that tapped breadth of vocabulary knowledge and one that tapped depth of vocabulary knowledge, in order to obtain a more complete assessment of this construct. The British Picture Vocabulary Scale–Third Edition (BPVS; Dunn et al., 2009) provided a measure of breadth of vocabulary knowledge. Each child was asked to select one of four pictures that best showed the meaning of a word spoken aloud by the experimenter. The test was administered and scored according to the guidelines in the manual.

[1] The stories were modified from materials constructed for an ongoing project conducted by the Language and Reading Research Consortium (2015). Minor modifications to vocabulary and grammar were made to make the stories suitable for speakers of British English.

Internal consistency was not reported in the manual. Therefore, Cronbach's alpha was calculated for this sample of children ($\alpha = .97$). The Word Associations subtest from the Clinical Evaluation of Language Fundamentals–Fourth Edition (CELF-IV; Semel, Wiig, & Secord, 2006) tapped semantic fluency and depth of vocabulary knowledge. In each trial, the child was asked to provide as many words as possible in 1 min from a specified category (items of clothing that people wear, animals, foods that people eat, or jobs that people do). The first category was a practice trial. The test was administered and scored according to the manual. We could not calculate internal consistency for this task, but the test–retest reliability value reported in the manual for this age range is good (.96–1.00).

Working memory measures

Each child completed four assessments of verbal working memory. There were two measures of the ability to store information (simple span) and two measures of the ability to store and process information (complex span). Internal consistency was not reported in the manual for the working memory measures, so it was calculated for the current sample.

Simple span measures. Each child completed the Word List Recall and Digit Recall subtests from the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001). In these tasks, children were asked to recall either lists of words or strings of digits spoken by the experimenter. The tasks were administered and raw scores were calculated according to the test manual. There were six trials at each level of difficulty. Once a child had correctly completed four of the trials on a level, the assessor moved on to the next level, as directed by the manual. The child received credit for any trials not completed on the previous level. The assessor stopped testing once three trials on a level were answered incorrectly. Raw scores (total number of trials correct, including those given credit for as a result of moving on to the next level) and standardized scores are presented in Results. Cronbach's alpha based on the current sample of children was good ($\alpha = .85$) for the word span and excellent ($\alpha = .90$) for the digit span.

Complex working memory span. Each child completed the Listening Recall and Counting Recall subtests from the WMTB-C (Pickering & Gathercole, 2001). In the Listening Recall task, children listened to short sentences, made a true/false judgment about each one, and then recalled the final word in each one. This test began with one sentence and increased in difficulty by the addition of sentences in each set (two sentences, three sentences, etc.). In the Counting Recall task, children counted the number of dots on a page and recalled the total. The test increased in difficulty in that children were required to count more than one pattern of dots over successive pages and then to recall the totals in the correct order. Both tests were administered according to the manual and followed the same progression and discontinuation rules as described above. Cronbach's alpha based on the current sample was good ($\alpha = .88$) for both the listening span and counting span ($\alpha = .90$).

Results

The results are reported in two sections. The first section concerns developmental comparisons of performance on all tasks. The second section concerns analyses that explore the interrelations between vocabulary, working memory, and inference performance. Before analysis, the "Q–Q plots" for all measures were examined separately for each year group. Q–Q plots provide a means of assessing deviation from a normal distribution (Cohen, Cohen, West, & Aiken, 2003). The plots indicated that the data were normally distributed. The data distributions were also checked for outliers. All analyses were run on the original dataset and also on a dataset that was adjusted to remove outliers following the recommendations of Tabachnick and Fidell (2007), whereby outlier data points are changed to the next highest/lowest (non-outlier) number. Less than 1% of datapoints were replaced in this way. The patterns of analyses were the same for both analyses, and so the analyses reported here were those conducted on the original data.

Table 2
Mean proportionate total correct scores (and standard deviations) for first responses and after prompts on the local and global coherence inference questions.

Inference type	6 years	8 years	10 years
Local			
First response	.49 (.16)	.59 (.19)	.67 (.16)
After prompts	.60 (.15)	.67 (.17)	.74 (.15)
Global			
First response	.67 (.14)	.72 (.17)	.80 (.13)
After prompts	.73 (.15)	.81(.16)	.89 (.10)

Developmental comparisons on all measures

First we report the performance on the inference task, followed by performance on the vocabulary and memory measures.

Performance on local and global inference questions

The proportion of total correct items (adjusted to take general knowledge into account, as described above) for local and global coherence inferences were the dependent variables in the analyses reported in this section. Before analysis, the data were arcsine transformed, as is recommended to stabilize variance and normalize proportional data (Sheskin, 2003). The pattern for the analyses of the raw scores was the same as that of the arcsine-transformed scores, so the analyses reported here are those conducted on the raw proportionate scores. Two analyses are reported below. In the first analysis, the dependent variable is performance when the question was first asked. In the second analysis, the dependent variable is performance that includes correct responses provided after a prompt. The mean proportions of correct responses on the local coherence and global coherence inference questions are shown in Table 2.

Performance on local and global inference questions: First responses. A mixed factor ANOVA with age (6, 8, or 10 years) as a between-participants factor and inference type (local or global) as a within-participants factor was conducted. There was a significant main effect of age, $F(2, 127) = 14.78$, $p < .001$, $\eta_p^2 = .19$. Post hoc Tukey tests ($p < .05$) indicated that the 6-year-olds obtained significantly lower scores than the 8- and 10-year-olds and that the 8-year-olds obtained significantly lower scores than the 10-year-olds (*M*s = .58, .65, and .74 in order of increasing ages). There was also a significant main effect of inference type, $F(1, 127) = 106.89$, $p < .001$, $\eta_p^2 = .46$, because the children performed best on the global inference questions ($M_{local} = .58$, $M_{global} = .73$). The interaction was not significant, $F(2, 127) < 1.0$, *ns.*

Performance on local and global inference questions: After prompts. As is evident from Table 2, performance improved with the use of prompts. The difference between unprompted and prompted scores (collapsed across age group) was significant for both types of inference: local, $t(129) = 15.72$; global, $t(129) = 12.08$; *ps* < .001. A mixed factor ANOVA with the same design as before was conducted on the scores that included prompted responses. The pattern of results was the same as for the analysis of unprompted scores, with main effects of age, $F(1, 127) = 14.21$, $p < .001$, $\eta_p^2 = .18$, and inference type, $F(1, 127) = 127.82$, $p < .001$, $\eta_p^2 = .50$, and no interaction between the factors, $F(2, 127) < 1.0$, *ns.*

Performance on vocabulary and working memory measures

The mean raw scores (total number of items correct for the vocabulary measures and total number of trials correct for the word, digit, listening, and counting working memory measures) for each age group are reported in Table 3. The mean standardized scores were available for all measures with the exception of the Word Associations task, for which standardized scores are not reported in the manual. The standardized scores were all within an age-appropriate range. Criterion reference scores were available for each age group on the Word Associations task, and all children met the minimum criterion score for their age. A small number of children ($n = 10$) had missing data due to absence at a testing session.

Table 3
Mean raw scores and standardized scores (with standard deviations) on the vocabulary and working memory measures

Measure	6 years	8 years	10 years	*F* value
BPVS				
Raw score (max = 168)	89.93[a] (10.74)	113.02[b] (14.80)	129.67[c] (11.30)	113.21***
Standardized score	97.68 (8.97)	102.86 (13.04)	95.02 (10.44)	
Word Associations				
Raw score	25.43[a] (6.18)	38.69[b] (7.36)	41.49[b] (7.60)	64.32***
Word List Recall				
Raw score (max = 42)	17.25[a] (2.93)	19.41[b] (2.90)	21.37[c] (4.40)	15.27***
Standardized score	104.56 (12.77)	101.62 (13.88)	103.10 (17.03)	
Digit Recall				
Raw score (max = 54)	27.09[a] (3.59)	30.34[b] (4.07)	32.88[c] (6.17)	15.72***
Standardized score	109.95(13.16)	108.95 (11.69)	112.53 (19.62)	
Listening Recall				
Raw score (max = 36)	5.68[a] (3.09)	9.72[b] (2.45)	11.79[c] (3.38)	46.66***
Standardized score	88.26 (16.95)	94.38 (11.75)	98.70 (17.58)	
Counting Recall				
Raw score (max = 42)	12.71[a] (4.15)	17.82[b](3.75)	19.78[b] (5.04)	28.54***
Standardized score	90.03(13.96)	87.14(13.65)	86.25 (16.69)	

Note. "BPVS" refers to the British Picture Vocabulary Scale–Third Edition (Dunn et al., 2009). Raw scores are presented on the upper row, and (where available) standardized scores are presented on the lower row. "Max" denotes maximum possible score. For raw scores, values with the same superscript are not significantly different from each other. Unless stated below, $N = 44$ for 6-year-olds, $N = 43$ for 8-year-olds, and $N = 43$ for 10-year-olds. Children were included wherever they had a test score. Missing data: BPVS (8 years, $n = 1$); Counting Recall (6 years, $n = 3$; 8 years, $n = 4$; 10 years, $n = 3$); Digit Recall (6 years, $n = 1$; 8 years, $n = 2$; 10 years, $n = 3$); and Word List Recall (8 years, $n = 1$).
*** $p < .001$.

A one-way ANOVA on the raw scores for each measure was conducted with age group as a between-participants factor. As predicted, there were age differences for all measures shown by significant *F* values (see Table 3; all *p*s < .001). Post hoc Tukey tests ($p < .05$) revealed that for most measures there were significant differences in performance between each successive age group in the following order: 6 < 8 < 10 years. The two exceptions were the Word Associations and Counting Recall tasks. Here, both the 8- and 10-year-olds obtained significantly higher scores than the 6-year-olds, but these two older groups did not differ from each other.

The 6-year-olds performed poorly on both complex span measures of working memory, with 27.3% of these children not passing the first level of the Listening Recall task and a further 61.4% not progressing beyond the one-span level. For the Counting Span task, 22.7% did not progress beyond the one-span level. This is in line with Gathercole, Pickering, Ambridge, and Wearing's (2004) observation that the task demands for these complex span measures are too high for very young children and that these complex span measures might not be sensitive to individual differences. On that basis, the complex span tasks were not included in further analyses for this age group.

Correlations between inference, vocabulary, and working memory

First, zero-order correlations were conducted for each age group separately to examine the interrelations between variables. Composite scores were produced to ensure a broad and comprehensive indicator of the vocabulary and memory constructs: vocabulary (BPVS and Word Associations), simple memory span (Word List Recall and Digit Recall), and complex memory span (Listening Recall and Counting Recall). The working memory composites support the distinction between tasks tapping short-term storage in the phonological loop and those tapping the central executive (Pickering & Gathercole, 2001). Reducing the number of measures to form composites also met the requirement of 10 data points per predictor for the multiple regression analyses reported next (as recommended by Tabachnick & Fidell, 2007). The scores on the first response to the inference questions were used. All scores were converted to *z* scores (for each age group separately) to ensure that measures were on a comparable scale (see Table 4).

Table 4
Correlations between local and global coherence inference performance and vocabulary, simple span, and complex span (excluding 6-year olds).

Age	Measure	Global inference	Vocabulary composite	Simple composite	Complex composite
6 years	Local	.40**	.56**	.44**	–
	Global		.60**	.31*	–
8 years	Local	.63***	.51**	.16	.25
	Global		.58**	.28	.37*
10 years	Local	.33*	.22	.20	.08
	Global		.67**	.37*	.39**

Note. $N = 44$ for 6-year-olds, $N = 43$ for 8-year-olds, and $N = 43$ for 10-year-olds.
* $p < .05$.
** $p < .01$.

A small number of children ($n = 10$) had missing data on one of the two subtests used to produce the composites. In these cases, the mean performance on the missing subtest for the relevant age group was used so that a composite could still be calculated for each of these children. This amounted to 18 data points in total (2.36% of all data points).[2]

Because of the small sample size, we discuss correlations both in terms of statistical significance and in relation to effect sizes where .10 represents a small effect, .30 a moderate effect, and .50 a large effect (Field, 2005). In all age groups, the two inference measures were significantly correlated and the effects were moderate to large. Of note, vocabulary was more strongly related to inference making than was memory. The vocabulary composite was significantly correlated with local coherence inference making for the 6- and 8-year-olds and with global coherence inference making for all age groups. All effect sizes were large. For the youngest age group, simple span was significantly correlated with local coherence inference making, with a moderate effect size, but a significant relation was not found for the older age groups. The simple and complex working memory measures (where used) were significantly correlated with global coherence inference making for all age groups, where the effects were moderate (with the exception of simple span for the 8-year-olds, where the effect was small and did not reach significance).

Do working memory and vocabulary each explain independent variance in inference making in different age groups?

A series of multiple regression analyses were conducted to determine the relative contributions of memory and vocabulary to the two types of inference. Fixed-order hierarchical multiple regression analyses were conducted for each inference type and each age group separately. Vocabulary, simple span, and complex span (8- and 10-year-olds only) were entered as predictors in the analyses.

Simple span was added at the first step, followed by complex span (8- and 10-year-olds only) and vocabulary. Previous research has found that vocabulary mediates the relationship between memory and inference generation (Chrysochoou et al., 2011); therefore, the influence of the memory measures before the inclusion of vocabulary could be determined using this order of variables. In addition, complex span tasks have been found to be more highly related to reading comprehension than simple span tasks (Daneman & Merikle, 1996). Therefore, this order of memory variables also ensured that any influence of simple span could be identified. Table 5 shows the results for the multiple regression analyses for the local coherence inferences.

For local coherence inferences, simple span and vocabulary explained variance in performance for the 6-year-olds, although (as indicated by the beta values) only vocabulary was a significant predictor when both variables were included in the model (see Table 5). For the 8-year-olds, only vocabulary significantly accounted for significant variance in local coherence inference making. In contrast, for the 10-year-olds, none of the measures significantly explained variance.

[2] A full correlation matrix for each age group is provided in the Appendix.

Table 5
Multiple regression with local coherence inference performance as the dependent variable and vocabulary, simple span, and complex span (excluding 6-year-olds) as predictors.

Age	Step	Variable	R^2	ΔR^2	Final β
6 years	1	Simple	.19	.19**	.25
	2	Vocabulary	.36	.17**	.46**
8 years	1	Simple	.03	.03	.02
	2	Complex	.08	.05	.03
	3	Vocabulary	.26	.18**	.49**
10 years	1	Simple	.04	.04	.15
	2	Complex	.04	.00	–.06
	3	Vocabulary	.07	.03	.18

Note. Standardized beta values are given for the final model with all predictors. "Simple" refers to simple span, and "complex" refers to complex span.
** $p < .01$.

Table 6
Multiple regression with global coherence inference performance as the dependent variable and vocabulary, simple span, and complex span (excluding 6-year-olds) as predictors.

Age	Step	Variable	R^2	ΔR^2	Final β
6 years	1	Simple	.10	.10*	.08
	2	Vocabulary	.37	.27***	.57***
8 years	1	Simple	.08	.08	.13
	2	Complex	.20	.12*	.15
	3	Vocabulary	.37	.17**	.48**
10 years	1	Simple	.13	.13*	.05
	2	Complex	.20	.07	.17
	3	Vocabulary	.48	.28***	.59***

Note. Standardized beta values are given for the final model with all predictors. "Simple" refers to simple span, and "complex" refers to complex span.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

For global coherence inferences, simple span and vocabulary explained additional variance in performance for the 6-year-olds, although (similar to the local coherence analyses) only vocabulary predicted unique variance when both variables were taken into account (see Table 6). For the 8-year-olds, complex span and vocabulary significantly accounted for additional variance in global coherence performance, although (again) only vocabulary was a significant predictor when all variables were taken into account. For the 10-year-olds, simple span and vocabulary explained additional variance in global coherence performance, but only vocabulary significantly predicted unique variance.

Does vocabulary mediate the influence of working memory on inference making in different age groups?

As outlined in the Introduction, the literature indicates that vocabulary may mediate the influence of memory on inference making. The data in Tables 5 and 6 demonstrate that both vocabulary and short-term memory made significant contributions to local inference making in the 6-year-olds and that vocabulary and memory (short-term span for 6- and 10-year-olds and complex span for 8-year-olds) predicted global inference making in all age groups. The analyses reported next assess whether or not vocabulary mediated the influence of memory on inference making.

The total effect of working memory on inference making is presented in Fig. 1, the mediation path assessed is presented in Fig. 2, and the labels used to denote each path are also referred to in

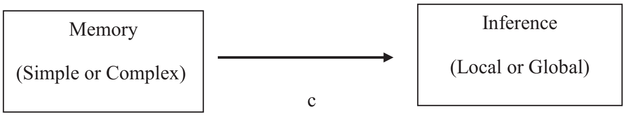

Fig. 1. Total effect of working memory on inference making.

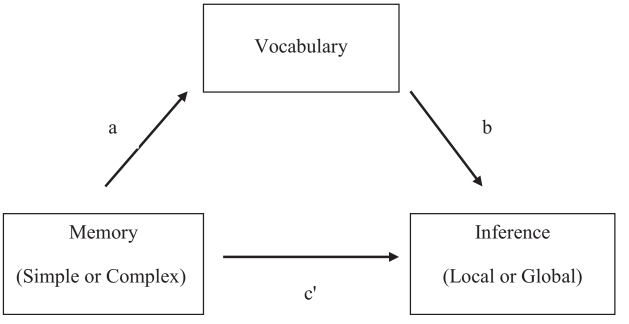

Fig. 2. Mediated effect of working memory on inference making.

Table 7
Mediated effect of working memory via vocabulary for local coherence inferences.

Age	Variable	Path	β	SE	t	PE	BC CI
6 years	Simple	a	.38	.13	2.87**	.21	.08–.44
		b	.57	.17	3.36**		
		c	.51	.16	3.15**		
		c′	.29	.16	1.85		

Note. "Simple" refers to simple span. "Path" refers to the paths shown in Figs. 1 and 2. "*β*" refers to beta coefficient, "*SE*" refers to standard error, "*t*" denotes *t*-test result, "*PE*" refers to point estimate, and "*BC CI*" refers to bias-corrected 95% confidence interval.
** $p < .01$.

Tables 7 and 8. A set of mediation analyses was conducted based on the causal steps strategy (see Baron & Kenny, 1986) and Preacher and Hayes's (2008) bootstrap (1000) resample procedure using an SPSS macro for this procedure, which gives point estimates (*PE*s) and bias-corrected confidence intervals (*BC CI*s) for the indirect effect. Where zero is not included in the 95% confidence interval for the mediated effect, the effect is statistically different from zero and mediation is confirmed.

For local coherence inferences, simple span predicted significant variance in local coherence inference making for the 6-year-olds (see Table 5). The mediation analyses demonstrated that vocabulary mediated its effect; the effect of working memory on inference making (c′ path) was not significant after taking vocabulary into account. Critically, mediation was confirmed because zero was not included in the 95% confidence interval for the mediated effect (see Table 7). Mediation analyses were not conducted for the 8- and 10-year-olds because the working memory measures did not predict unique variance in local coherence scores (see Table 5).

For global coherence inferences, vocabulary mediated the relationship between working memory and inference for all age groups. As shown in Table 6, although simple span predicted inference for the 6- and 10-year-olds, the effects of working memory on inference were mediated by vocabulary at both ages because the path coefficient for the regression of this working memory measure on inference was no longer significant when vocabulary was entered into the equation (c′ path). Mediation was confirmed because zero was not included in the 95% confidence interval for the mediated effect (see Table 8). For the 8-year-olds, complex span was related to global coherence inference making (see Table 6).

Table 8
Mediated effect of working memory via vocabulary for global coherence inferences.

Age	Variable	Path	β	SE	t	PE	BC CI
6 years	Simple	a	.38	.13	2.87**	.27	.10–.57
		b	.71	.17	4.17***		
		c	.36	.17	2.12*		
		c′	.09	.16	0.59		
8 years	Complex	a	.21	.07	3.17**	.14	.04–.28
		b	.65	.18	3.67***		
		c	.22	.09	2.54*		
		c′	.08	.08	0.98		
10 years	Simple	a	.37	.13	2.84**	.28	.07–.57
		b	.77	.16	4.92***		
		c	.41	.16	2.52*		
		c′	.13	.14	0.89		

Note. "Simple" refers to simple span, and "complex" refers to complex span. "Path" refers to the paths shown in Figs. 1 and 2. "*β*" refers to beta coefficient, "*SE*" refers to standard error, "*t*" denotes *t*-test result, "*PE*" refers to point estimate, and "*BC CI*" refers to bias-corrected 95% confidence interval.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

However, as for the analysis for the two other age groups above, vocabulary mediated this effect because the path coefficient for the regression of complex span on inference was no longer significant when vocabulary was entered into the equation (c′ path). The mediated effect of complex span via vocabulary was significant because zero was not included in the 95% confidence interval (see Table 8).

Discussion

The main aims of this study were to explore developmental differences in local and global coherence inference making and to determine the extent to which vocabulary and working memory influenced performance. In line with expectations and previous research, the 6-year-olds were able to generate these inferences from short narrative text, but significant improvements were found with age. Critically, our work indicates that vocabulary and working memory differed in the extent to which they predicted performance on both types of inference between age groups. Specifically, there was evidence that vocabulary knowledge was critical both independently and through its role in supporting the memory processes required for global inference making in particular.

First, let us consider performance on the local coherence inferences. In line with previous research, performance improved with age (Ackerman, 1986). To integrate information between sentences in order to establish local coherence, children needed to understand and make use of pronouns and also synonyms. The previous literature on pronoun and synonym comprehension has focused on individual differences in relation to measures of reading comprehension skill in general. This body of work finds that children who differ in reading comprehension skill differ in their understanding and use of pronouns and synonyms to integrate different propositions in a text (Ehrlich & Remond, 1997; Megherbi & Ehrlich, 2005; Oakhill, 1983; Oakhill & Yuill, 1986; see also Cain & Nash, 2011, for developmental improvements in knowledge of other cohesive devices). We propose that poorer use of these signaling cues may have limited local coherence inference making for the children in our study and could, in part, explain the developmental differences.

The source of the poorer use of these signals appears to be different for the 6-, 8-, and 10-year-olds. Vocabulary was a particularly important predictor of local coherence inference making for the two youngest age groups, a finding supported by previous research on inference making (Chrysochoou & Bablekou, 2010; Chrysochoou et al., 2011) and reading comprehension in general (Seigneuric & Ehrlich, 2005). The current study is the first to show that vocabulary knowledge is specifically related to 6- to 10-year-olds' ability to make inferences to establish local coherence. We cannot determine whether children lacked the category-relevant knowledge, failed to activate this, and/or failed to use this information during text processing from our data. However, we believe that a lack of relevant knowledge per se is an unlikely source of these developmental differences because few children failed the background knowledge check questions for the global inference measures. Thus, we propose that speed of knowledge retrieval and integration during text presentation are more likely sources of failure to establish local coherence (as proposed by Perfetti et al., 2008, for adults). We are currently exploring online text processing in these age groups to examine these possibilities.

Although the zero-order correlations and regression analyses indicated that working memory played a role in the determination of the youngest children's local coherence inference performance, working memory was not a unique predictor after taking vocabulary knowledge into account. The youngest children's poorer memory skills may have limited their ability to retain the information from two successive sentences in order to integrate their meanings when constructing their situation model of the text. However, the mediation analysis confirmed that for the 6-year-olds the influence of memory was fully mediated by vocabulary, similar to the findings in other work with older children (Chrysochoou et al., 2011). Children with richer vocabulary knowledge will be better able to retain information in verbal working memory (Cain, 2006; Nation et al., 1999), which supports our finding of a mediation effect. A complex interaction between vocabulary and verbal working memory may underpin developmental differences in this type of inference making, and further work is needed to test the proposal that speed of knowledge retrieval and/or maintenance of accurate representations of meaning were the source of local coherence inference difficulties.

In contrast to the pattern of prediction for the youngest age groups, neither vocabulary nor working memory explained performance on the local coherence inferences for the oldest age group. The older children had superior memory and vocabulary skills, which may have been sufficient for them to establish local coherence. However, the 10-year-olds were not at ceiling on the task, so this explanation is unsatisfactory. One possibility is that they had sufficient vocabulary and working memory skills to perform the task, but additional gains on these questions could come from strategic processing. For example, older children may have greater awareness of which information in a text is critical for the generation of local and global inferences. We are currently investigating developmental differences in children's use of text information in the generation of both inference types to explore this issue further. Work with poor comprehenders (Yuill & Oakhill, 1988) also shows that children can be taught how to make inferences and that this boosts both inference making skill and reading comprehension.

We now turn to a discussion of the findings on global coherence inferences. As predicted, and in line with other work, there were developmental improvements on the global coherence inferences (Chrysochoou & Bablekou, 2010; Schmidt & Paris, 1983). It has been suggested that younger children tend to view statements in a text as independent pieces of information and do not always link successive ideas together, particularly if they are not presented in succession (Schmidt & Paris, 1983). The local coherence inferences were signaled, indicating when successive pieces of text should be linked. In contrast, the global coherence inferences were not signaled and typically required the listener to link information across a number of different nonconsecutive sentences. Thus, one reason for failing to generate these inferences is a piecemeal processing style. Another reason is the processing demands of this type of inference. The global coherence inferences required the integration of information presented in different nonconsecutive sentences. The association between working memory and performance on this type of inference probably arose because of this demand on memory.

Age differences in global coherence inference ability persisted even when differences in the necessary background knowledge were controlled (indeed, there were only three instances where a child failed to correctly answer a global coherence inference and lacked the necessary knowledge). Thus, this study adds to a growing body of work demonstrating that the ability to make coherence inferences is not solely dependent on having the relevant background knowledge (Barnes et al., 1996; Cain & Oakhill, 1999; Cain, Oakhill, Barnes, & Bryant, 2001). The findings support the idea that children need to know when and how to draw on background knowledge during text comprehension and that this ability might improve with age (Cain & Oakhill, 1999).

Our work advances our understanding of the development of inference making by demonstrating that at all three ages, vocabulary was the most important factor explaining performance on the global coherence measures and vocabulary fully mediated the relationship between global coherence inference making and working memory. This finding is supported by previous work assessing this inference type with older children (Chrysochoou et al., 2011). Thus, although children demonstrated relevant background knowledge when assessed, they did not always use this to generate or encode the inference. Furthermore, the information to be linked was typically presented in words that were semantically associated; for example, "pet," "furry," "playful," and "kennel" were all used to indicate that the story was about a dog. It may be that, for some children, the associations between these links were not sufficiently strong to activate the concept to which they were all related. The developmental improvements found in vocabulary knowledge and the richness of semantic networks, in which the associations between co-occurring concepts are encoded (Metzger et al., 2008), may assist in inference making that relies on background knowledge. Future work, using online measures to assess whether children use successive pieces of information, may help to further our understanding of the production of this inference type.

In this study, children made a greater number of global coherence inferences than local coherence inferences, in contrast to other work that has compared the generation of these inference types during reading (Cain & Oakhill, 1999). Although the questions tap different elements of the story and, therefore, are not directly comparable, this finding requires consideration. One explanation could be that the global coherence inferences (particularly those that were thematic in nature, e.g., the setting of a story or a character's identity) were more foregrounded because the text made a number of references to these elements. Support for this explanation comes from work on the centrality effect, which finds that information that is more central to the overall story meaning is more likely to be remembered than information that is peripheral to overall story meaning (e.g., see Albrecht & O'Brien, 1993; Miller & Keenan, 2009). Future work comparing these inference types should consider this difference.

There are several educational implications that stem from this work. First, performance for all age groups on both types of inference improved after prompting underspecified responses. This finding identifies prompting as a useful means to demonstrate a child's potential (or competence) and also as a tool to indicate the standard of coherence required of comprehension (van den Broek, Risden, & Husebye-Hartman, 1995). Second, the finding that all age groups made both global and local coherence inferences indicates that children of this age engage in the processing required to construct a representation of the entire text rather than engaging in piecemeal line-by-line processing. Third, the importance of vocabulary knowledge to inference making and the way in which vocabulary mediated the relations between working memory and inference making supports the critical role of background knowledge in text comprehension (Eason, Goldberg, Young, Geist, & Cutting, 2012). It also lends weight to other research proposing that children need to be taught how to use both the text and background knowledge as sources of information to guide inference generation and text comprehension (Brandão & Oakhill, 2005; Raphael & Wonnacott, 1985).

There are several limitations to this work that need to be addressed in future work in order to understand fully the roles of vocabulary and working memory in children's inference making. First, although our sample size had adequate statistical power to detect the contributions of key skills, replication with a larger sample size is required to confirm the developmental differences. Second, we did not include independent assessments of strategic knowledge, so we do not know to what extent different age groups were aware of different sources of knowledge and strategies that may influence language comprehension (Brandão & Oakhill, 2005). Third, although we drew on a large literature indicating that the same processes underpin the comprehension of written and spoken text, we note the evidence for a greater influence of attention on listening comprehension (Cain & Bignell, 2014). Future work should consider the extent to which this might specifically affect inference making. Finally, all of our measures were offline. Language comprehension is a dynamic process. Thus, it is essential to investigate how both vocabulary and memory (both short-term and working memory) interact during online language processing to identify when and where they influence the inference-making process. We are pursuing this line of inquiry in ongoing work.

In summary, few studies have experimentally contrasted children's ability to generate different types of inference, and none has developmentally explored the unique roles of vocabulary and working memory in local and global inference making. This study has highlighted different roles for vocabulary and working memory in relation to these two types of inference at different ages. This not only adds to our understanding of inference development but also indicates factors that may limit language comprehension in the classroom.

Acknowledgments

This research was conducted at Lancaster University and was funded by an Economic and Social Research Council (ESRC) PhD studentship awarded to the first author. We thank all of the teachers and pupils at the participating schools for their help with this research.

Appendix

Table A1. Correlations between local and global coherence inference performance and the vocabulary and working memory measures for the 6-year-olds.

Measure	1	2	3	4	5	6	7	8
1. Local	1.00	.40**	.47**	.43**	.43**	.33*	.39**	–.04
2. Global			.56***	·40**	.20	.34*	.24	–.02
3. BPVS		1.00	1.00	.29[a]	.35*	.35*	.52***	.29[b]
4. Word Associations				1.00	.18	.25	.44**	.26
5. Word List Recall					1.00	.49**	.44**	.28
6. Digit Recall						1.00	.42**	.37*
7. Listening Recall							1.00	.40**
8. Counting Recall								1.00

Note. "Local" refers to local coherence, and "global" refers to global coherence. "BPVS" refers to the British Picture Vocabulary Scale–Third Edition (Dunn et al., 2009). Unless stated below, $N = 44$. Children were included wherever they had a test score. Missing data: Counting Recall ($n = 3$) and Digit Recall ($n = 1$).

[a] $p = .052$.
[b] $p = .063$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table A2. Correlations between local and global coherence inference performance and the vocabulary and working memory measures for the 8-year-olds.

Measure	1	2	3	4	5	6	7	8
1. Local	1.00	.63***	.45**	.37*	.27	.01	.24	.18
2. Global			.61***	·34*	.45**	.05	.30[a]	.33*
3. BPVS		1.00	1.00	.30[b]	.43**	.03	.36*	.38*
4. Word Associations				1.00	.28	.08	.38*	.09
5. Word List Recall					1.00	.56***	.22	.06
6. Digit Recall						1.00	.01	–.08
7. Listening Recall							1.00	.47**
8. Counting Recall								1.00

Note. "Local" refers to local coherence, and "global" refers to global coherence. "BPVS" refers to the British Picture Vocabulary Scale–Third Edition (Dunn et al., 2009). Unless stated below, $N = 43$. Children were included wherever they had a test score. Missing data: BPVS ($n = 1$), Counting Recall ($n = 4$), Digit Recall ($n = 2$), and Word List Recall ($n = 1$).

[a] $p = .051$.
[b] $p = .056$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table A3. Correlations between local and global coherence inference performance and the vocabulary and working memory measures for the 10-year-olds.

Measure	1	2	3	4	5	6	7	8
1. Local	1.00	.33*	.35*	.01	.23	.14	.11	.02
2. Global		1.00	.59***	.49**	.34*	.34*	.43**	.22
3. BPVS			1.00	.30*	.31*	.26	.42**	.12
4. Word Associations				1.00	.34*	.28	.33*	.05
5. Word List Recall					1.00	.61***	.19	.35*
6. Digit Recall						1.00	.32*	.48**
7. Listening Recall							1.00	.34*
8. Counting Recall								1.00

Note. "Local" refers to local coherence, and "global" refers to global coherence. "BPVS" refers to the British Picture Vocabulary Scale–Third Edition (Dunn et al., 2009). Unless stated below, $N = 43$. Children were included wherever they had a test score. Missing data: Counting Recall ($n = 3$) and Digit Recall ($n = 3$).

* $p < .05$.

** $p < .01$.

*** $p < .001$.

References

Ackerman, B. P. (1986). Referential and causal coherence in the story comprehension of children and adults. *Journal of Experimental Child Psychology, 41*, 336–366.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology, 19*, 1061–1070.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.). *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

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

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An online assessment of causal reasoning during comprehension. *Memory & Cognition, 18*, 65–71.

Brandão, A. C. P., & Oakhill, J. (2005). "How do you know this answer?" Children's use of text data and general knowledge in story comprehension. *Reading and Writing, 18*, 687–713.

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory, 14*, 553–569.

Cain, K., & Bignell, S. (2014). Reading and listening comprehension and their relation to inattention and hyperactivity. *British Journal of Educational Psychology, 84*, 108–124.

Cain, K., & Nash, H. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology, 103*, 429–441.

Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure. *Reading and Writing, 11*, 489–503.

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

Cain, K., Oakhill, J., Barnes, M., & Bryant, P. (2001). Comprehension skill, inference making ability, and their relation to knowledge. *Memory & Cognition, 26*, 850–859.

Cain, K., Oakhill, J., & Bryant, P. E. (2000). Investigating the causes of reading comprehension failure: The comprehension–age match design. *Reading and Writing, 12*, 31–40.

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

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.

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

Chrysochoou, E., & Bablekou, Z. (2010). Phonological loop and central executive contributions to oral comprehension skills of 5.5 to 9.5 year old children. *Applied Cognitive Psychology, 25*, 576–583.

Chrysochoou, E., Bablekou, Z., & Tsigilis, N. (2011). Working memory contributions to reading comprehension components in middle childhood children. *American Journal of Psychology, 124*, 275–289.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.
Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*, 465–472.
Dunn, L. M., Dunn, D., Sewell, J., Styles, B., Brzyska, B., Shamsan, Y., et al (2009). *The British Picture Vocabulary Scale* (3rd ed.). London: GL Assessment.
Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology, 104*, 515–528.
Ehrlich, M., & Remond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. *British Journal of Developmental Psychology, 15*, 291–309.
Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
Florit, E., Roch, M., Altoe, G., & Levorato, C. M. (2009). Listening comprehension in preschoolers: The role of memory. *British Journal of Developmental Psychology, 27*, 935–951.
Florit, E., Roch, M., & Levorato, C. M. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes, 48*, 119–138.
Gathercole, S. E., Pickering, S., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177–190.
Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix providing multilevel analyses of text characteristics. *Educational Researcher, 40*, 223–234.
Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.
Hannon, B., & Frias, S. (2012). A new measure for assessing the contributions of higher level processes to language comprehension performance in preschoolers. *Journal of Educational Psychology, 104*, 897–921.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*, 127–160.
Kendeou, P., van den Broek, P., White, M., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*, 765–778.
Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95*, 163–182.
Language and Reading Research Consortium. (2015). *The dimensionality of language ability in young children.* submitted for publication.
Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology, 58*, 88–111.
Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1424–1429.
Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology, 29*, 327–365.
McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440–466.
Megherbi, H., & Ehrlich, M. (2005). Language impairment in less skilled comprehenders: The on-line processing of anaphoric pronouns in a listening situation. *Reading and Writing, 18*, 715–753.
Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., Reed, A. W., & Williams, D. (2008). Do children "DRM" like adults? False memory production in children. *Developmental Psychology, 44*, 169–181.
Miller, A. C., & Keenan, J. M. (2009). How word decoding skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of Dyslexia, 59*, 99–113.
Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139–158.
Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5–20.
Oakhill, J. (1983). Instantiation in skilled and less skilled comprehenders. *Quarterly Journal of Experimental Psychology A, 35*, 441–450.
Oakhill, J., & Yuill, N. (1986). Pronoun resolution in skilled and less skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech, 29*, 25–37.
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.
Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension skill and word to text integration processes. *Applied Cognitive Psychology, 22*, 303–318.
Pickering, S., & Gathercole, S. (2001). *Working Memory Test Battery for Children*. London: Psychological Corporation.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.
Raphael, T. E., & Wonnacott, C. A. (1985). Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly, 20*, 282–296.
Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes, 29*, 37–60.
Schmidt, C. R., & Paris, S. G. (1983). Children's use of successive clues to generate and monitor inferences. *Child Development, 54*, 742–759.
Seigneuric, A., & Ehrlich, M.-F. (2005). Contribution of working memory capacity to children's reading comprehension: A longitudinal investigation. *Reading and Writing, 18*, 617–656.
Semel, E., Wiig, E. H., & Secord, W. A. (2006). *Clinical evaluation of language fundamentals* (4th ed.). London: Harcourt Assessment.
Sheskin, D. J. (2003). *Handbook of parametric and nonparametric statistical procedures* (3rd ed.). Boca Raton, FL: CRC Press.
Stein, N. L., & Glenn, C. G. (1982). Children's concept of time: The development of story schema. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 255–282). New York: Academic Press.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing, 4*, 245–256.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.

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.

van den Broek, P. W., Risden, K., & Husebye-Hartman, 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.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1256–1271.

Yuill, N., & Oakhill, J. (1988). Effects of inference awareness training on poor reading comprehension. *Applied Cognitive Psychology, 2*, 33–45.