Four decades of research into children's reading comprehension: a personal review

Article (Accepted Version)

Four decades of research into children's reading comprehension: A personal review.

Jane Oakhill

University of Sussex, UK

Abstract

A substantial amount of research has focused on children's reading development and reading problems but, in comparison, there has been relatively little research into children's reading comprehension. This paper will provide an overview of the research that has investigated the skills and cognitive processes that support children's understanding of text, and will reflect on the implications of the findings presented in helping children to develop and improve their comprehension skills. The paper will conclude by considering which avenues of investigation still need further exploration.

Introduction

The fundamental goal of learning to read is to be able to understand text – whether it be for enjoyment, for some practical purpose, to learn or to find information. But comprehension is a complex construct, and it is difficult to know how it might be taught, or improved, without giving attention to the processes and skills that need to be orchestrated during the comprehension process. In this paper, I will outline the research that I have been carrying out, on and off, for about 40 years, in which I have endeavoured to investigate those underlying processes that support the end product (comprehension). The ultimate goal of this research has always been to discover which processes are less well-developed in poor comprehenders and which, thus, might benefit from explicit teaching and support in order to improve comprehension.

My interest in this area began in the late 1970s, when I observed that some children had good or very good word reading but, when asked about the story they had been reading, seemed to have little idea.[1] Why was it that some children could 'read' so well, yet comprehend so poorly? At the time there was almost no research into this question: The assumption seemed to be that, once children could decode the words on the page, then their language comprehension skills would come into play, and they would understand the text. That assumption was based on the premise that the children were all competent users of their native language. This was sometimes called "the decoding sufficiency" viewpoint, which argues that slow word decoding will detract from reading comprehension, with the

[1] These children were subsequently termed "less-skilled comprehenders" or "poor comprehenders" or, more accurately, children with a specific reading comprehension difficulty (SRCD). Throughout this paper, I will usually use the term "poor comprehenders", for brevity.

consequence that fast and automated decoding will be sufficient for reading comprehension. Thus, if children had problems with comprehension (in the absence of diagnosed language problems), that was because they had difficulties with word decoding.

At that time, there were many theories of word reading (e.g., see Singer & Ruddell, 1976) but little attention was given to reading comprehension, especially in children. In reading comprehension research in adults, it was generally agreed that the end product of comprehension was an integrated model of the text, but there was little research into the processes that contribute to that text representation. Furthermore, researchers did not apply ideas from the adult literature to the exploration of individual differences in children's reading comprehension. Thus, a typical perspective was that poor or slow decoding was a kind of bottleneck to reading comprehension (a version of the decoding-sufficiency theory). Thus, my research started with a more thorough comparison of word-recognition skill in skilled and less-skilled comprehenders.

Early ideas about specific reading comprehension difficulties

In the late 1970s, when I began this research, the only investigators who had explored the issue of comprehension problems in the absence of word recognition problems were Ward Cromer and colleagues (e.g., Cromer, 1970; Oakan, Wiener & Cromer, 1971; Steiner, Wiener & Cromer, 1971). Cromer identified two groups of poor comprehenders (both in Grade 5 and at college level). One group, who had poor vocabulary and decoding skills, as well as poor comprehension skills, he termed the 'deficit' group. The 'difference' group, on the other hand, had a comprehension problem, even though their vocabulary and decoding skills were commensurate with those of good comprehenders of the same age. Thus, the 'deficit' group had a general reading problem, whereas the 'difference' group had a specific comprehension problem. In this respect, the subjects in Cromer's difference group were very similar to the poor comprehenders mentioned above. But Cromer and colleagues were in a minority: The view that children could be truly good decoders but poor comprehenders was given little credibility.

Early studies of word and sentence-level skills

In the late 1970s and into the 1980s, Perfetti's Bottleneck Theory (e.g. Perfetti, 1985) was current as an account of reading comprehension problems. The idea was that not only a certain level of word reading ability, but also word-decoding speed and automaticity were crucial to avoid a bottleneck in processing, which would limit comprehension. Thus, it was possible that, although poor comprehenders were all good word readers, there might be more subtle differences in their word decoding. This idea provided the basis for several preliminary experiments. It was clear from Cromer's studies that the way in which such children were selected was critical, and that both single-word decoding skill and language comprehension needed to be taken into account when assessing children's reading comprehension. In most of my early studies, outlined below, reading ability was assessed using one or other version of *The Neale Analysis of Reading Ability* (e.g. 1997), precisely because it provides separate norms for word reading and reading comprehension. Importantly, this assessment is administered one-to-one, and any words that the child cannot read, or misreads, are supplied. Thus, if a child performs poorly on the comprehension questions at the end of a passage, that cannot be simply because they were unable to read the words in that passage. Teachers tended to be of the opinion that the poor comprehenders' problem was that they could read (that is, decode) the words, but didn't know their meanings, so good and poor comprehenders were matched for vocabulary knowledge, using the British Picture Vocabulary Scale (BPVS: Dunn, Whetton & Burley, 1997; a British version of the Peabody), or an adaptation of the Gates-MacGinite Vocabulary Subscale (MacGinitie, MacGinitie, Maria & Dreyer, 2000). Typical groups of good and poor comprehenders are shown in Table 1. It was not until quite recently that researchers have revisited the idea that representations of meaning at the word level (i.e. vocabulary) might be an interesting source of differences between good and poor comprehenders (see the later discussion of breadth vs. depth of vocabulary).

In an initial study, I assessed speed of decoding (for both words and non-words) and found no differences between the groups. A further study explored the speed and automaticity of access to the meanings of written words, using a picture/word interference task (see Figure 1). This Stroop-like task (in which children are required to name pictures, while ignoring words superimposed on the pictures) was adapted from a study by Golinkoff, Rosinski and Kukish (1975). They had used this task to investigate the relations between decoding, semantic access and comprehension. Their results showed that semantic access is no more automatic in good than in poor comprehenders. But in their experiment, as in almost all studies at that time, reading comprehension was limited by word recognition skill, so the implications of the results for children with specific reading comprehension difficulties are not clear.

As in Golinkoff et al.'s study, children were required children to name, as rapidly as possible, simple words and non-words (CVC trigrams, such as *pog, tiv),* or to name pictures of the objects, with or without words or non-words superimposed on them. The principle is the same as that that produces the Stroop effect: If word reading is automated and cannot be suppressed, then picture naming will be made more difficult if the name of a different object is superimposed on the picture (e.g. a picture of a cat, with the word *pig* superimposed). Thus, the difference in time taken to name pictures when compatible vs. conflicting words are superimposed on them gives a measure of the interference produced by conflicting words, and thus, the extent to which their meaning is automatically accessed. The children were required to name a set of pictures, as quickly as possible, in three different conditions: With congruent label superimposed, an incongruent label superimposed (as shown in Figure 1), or with a non-word superimposed. If any interference effect is semantic, rather than a general interference effect, it should be larger when conflicting labels are superimposed on the pictures than when nonwords are superimposed. In each case, the children's task was always to name the pictures, and ignore the words. There were two further control tasks, which measured the speed at which the words and non-words (trigrams) could be decoded in isolation, and a picture-naming task to ensure that subjects were able to name all the pictures accurately (for more details of this study, see Yuill & Oakhill, 1991).

There were no significant differences between the groups in accuracy or speed of word or trigram decoding and, although the good comprehenders were able to complete the picture-naming task more quickly than the poorer ones, both groups showed a semantic interference effect of a similar size. There were very few picture naming errors, and the only significant difference that emerged between the good and poor comprehenders was the difference in speed of picture naming in general. This result proved interesting in the light of much more recent studies, but at the time it did not catch my attention as it was not an expected finding. Indeed, it conflicted with other results: Perfetti, Finger and Hogaboam (1978), found no differences between good and poor comprehenders in the speed with which they could name colours, numerals and pictures.

In sum, the Bottleneck Hypothesis could not provide an adequate account of children's reading comprehension problems (providing they were matched for word recognition accuracy). However, there were already some indications that these children might have other issues with the representation of lexical information, to which I will return later in this paper.

Early studies also failed to find differences in syntactic skills (as assessed by the Test for Reception of Grammar: TROG2; Bishop, 2003). A full account of these studies of word and sentence-level processing and short-term memory can be found in Yuill and Oakhill (1991*).*

Thus far, the outcomes of these studies could simply be regarded as failures to reject the null hypothesis. Since they were conducted long before the days of “Open Science” it was impossible to pre-register and publish such studies. However, what could be concluded was that, if there were differences in automaticity of decoding between good and poor comprehenders who were matched for word reading accuracy, they were likely small and elusive. Thus, the results of those early studies caused me to stop and re-evaluate potential causes of comprehension problems, and it seemed more fruitful to move on to other lines of enquiry.

Inference skills (and memory ability)

In the 1970s, John Bransford and his colleagues were beginning to challenge the view that the representation of a text is a linguistic representation of the meaning of the sentences in that text. In a number of ingenious experiments, Bransford and colleagues showed that text comprehension is a “constructive and integrative process”: that readers go beyond the words and sentences to construct meaning.[2]

Research on developmental differences in inference skills by Paris and colleagues (e.g., Paris & Carter, 1973) raised the question as to whether, like younger vs. older children,

---

[2] With the inspiration and support of Phil Johnson-Laird (my DPhil supervisor), I began to consider children’s reading comprehension from a rather different perspective. Johnson-Laird’s research interests encompassed both text comprehension and deductive reasoning, but in both cases, his quest was to detail the processes that occur as people reason and construct meaning. It is interesting to note that the collection of papers in which he and Peter Wason republished an edited version of Bransford and McCarrell’s (1975) article is called: *Thinking: Readings in Cognitive Science*: i.e. in a book about thinking, rather than reading.

good vs. poor comprehenders might differ in their ability to infer information from text. Paris and colleagues used a false memory paradigm, where they asked children to remember a list of sets of three sentences, and then asked them whether or not they had heard specific sentences. For each original set, the children were presented with two sentences that they had actually heard, one sentence that was a valid inference from the information they had been presented, and one that was not. The good comprehenders more often correctly accepted the original, and rejected the invalid inference, sentences but they also made significantly more errors of a particularly interesting sort: they were more likely to say that they had *heard/read* the valid inference sentences than were the poor comprehenders (Oakhill, 1982). A later replication showed that this result held for both reading and listening (Oakhill, Yuill & Parkin, 1986). These results were interpreted as showing that the good comprehenders were making more text-connecting inferences than the poor comprehenders: a finding that has been confirmed in many different contexts since (see below).

Another study of inference in children (Oakhill, 1983), used a task that required instantiation of general terms (inspired by work in adults by Richard Anderson and colleagues: Anderson, Pichert, Goetz, Shallert, Stevens, & Trollip, 1976). The results showed that, given a list of sentences to remember, such as *The fish attacked the swimmer*, the skilled, but not the poor, comprehenders found specific instantiations of words that had not occurred in the sentences to be remembered (*shark*, in the example above), better clues to memory than the words that actually occurred (*fish*); an effect that was not present in the less-skilled group. This finding, although ostensibly on inference from context, can also be considered an instance of evidence that better comprehenders are making more context-relevant semantic associations from words, and presaged the studies of depth of vocabulary that came many years later.

Memory capacity

Another possibility was that some children simply had more general cognitive difficulties that manifested themselves as comprehension problems; for instance, remembering what they had just read. Indeed, in Oakhill's (1982) study of inference skill, the poor comprehenders showed a tendency to remember less information overall. Thus, perhaps their problems with comprehension were primarily a memory problem? This hypothesis was rendered more plausible by the fact that the Neale Analysis, as administered at the time, depended on memory: the children were not allowed to look back to the text, but had to respond to questions from memory.

Unlike typical dyslexic readers, poor comprehenders were not found to have any problems with short-term memory tasks, such as digit- or word-span (Yuill & Oakhill, 1991). Further investigation provided no evidence for a general deficit in memory for text in poor comprehenders. For instance, Oakhill (1984), asked groups of good and poor comprehenders to answer both literal and inference questions about short texts, in the first instance from memory, and then again with the text to refer to so that they could check their answers, or consult the text in the case of questions they had not been able to answer. The results showed that the poor comprehenders were worse both without the text, and with the text, even though they were encouraged to go back to the text to find or check their answers, and that in both conditions they had particular problems with inference questions. The finding that these problems persisted even when the text was available for them to refer to, clearly demonstrates that their problems could not simply be attributed to poor memory for the texts they had read. In a later study, Kate Cain and I[3] (Cain & Oakhill, 1999), conducted a study where we not only asked questions again with the text available, but also directed the children back to

[3] After I had completed my DPhil thesis, most of my studies were conducted in collaboration. In this area, my two primary collaborators (both of whom worked as postdoctoral fellows with me) were Kate Cain and Nicola Yuill and their contributions are evident in many of the studies described in this paper.

relevant parts of the text and asked the questions again. Both groups improved with the direction to the text, but even so, the poor comprehenders still performed more poorly than the comparison group on global coherence inferences.

Working memory and text comprehension

As mentioned above, poor comprehension in children who were good word readers could not be attributed to poor short-term memory. But in the early 1980s there was substantial interest in the relation between working memory and text comprehension, mainly driven by Daneman and Carpenter's (1990; 1983) work, and this seemed like a promising route to understanding the basis of text-comprehension problems in children.

My own studies showed that, while good and poor comprehenders showed no differences on classic short-term memory tasks (Oakhill, Yuill, & Parkin, 1986), they did show consistent differences on working memory tasks (including numerical WM tasks) (Yuill, Oakhill & Parkin, 1989). The groups did not differ on the least demanding versions of the tasks (processing and remembering two sets of words or digits) but as the working memory load increased, so did the difference between the groups. The largest group differences were found on a sentence-span task (perhaps because the sentence comprehension component of the task shares a lot of variance with text comprehension). However, no differences were obtained on a spatial working memory span task (Oakhill, Yuill & Garnham, 2011). Others have replicated and extended these findings: Children with poor comprehension skills tend to have lower working memory capacity than children with good comprehension (Cain, 2006; Nation, Adams, Bowyer-Crane, & Snowling, 1999). Even from on these initial findings, it was an obvious hypothesis that verbal working memory capacity might be an important cognitive ability that was limiting reading comprehension in some children because inference and integration of text, and other skills that are related to reading comprehension such as comprehension monitoring, typically require integration of information across several sentences in the text, and such integration is likely to be limited by working memory. Some early empirical work provided support for this hypothesis: children with poorer reading comprehension had greater difficulties across a range of tasks that required integration across the text. As an example, on a comprehension monitoring task[4], poor comprehenders only had greater difficulty in spotting and resolving inconsistencies in the text when the inconsistencies and the information that would help resolve them were separated by several sentences (Yuill, Oakhill & Parkin, 1989). When the inconsistent sentences were adjacent to the resolving information, the good and poor comprehenders showed a very similar level of performance. An example of inconsistent sentences is:

Tommy didn't share his sweets with his little brother.

When their mother saw this, she was very pleased with Tommy.

The apparent anomaly was resolved by using the further information that:

Tommy's brother was very fat and was on a diet.

which was presented either between the inconsistent sentence, or at the beginning of the text. None of the children had problems with the memory or norm questions (e.g., for the example above):

Was Tommy praised or blamed for what he had done?

Would someone usually be praised or blamed for not sharing their sweets?

Furthermore, independent measures of working memory have been found to predict inference making skill in typically developing readers (Cain, Oakhill, & Bryant, 2004). For children with poor comprehension skills, memory is particularly predictive of their

[4] The term "comprehension monitoring" is used in slightly different, though related, ways. It can be used to refer to a reader's judgement of their own understanding, but is more typically used to refer to a reader's realisation that something is amiss with their comprehension, and the corollary of that realisation: that some sort of comprehension repair Is needed. Error detection tasks are typically used to measure comprehension monitoring in children.

performance when the inference required the integration of information in sentences separated by several additional sentences (Barnes, Faulkner, Wilkinson, & Dennis, 2004; Cain, Oakhill, & Lemmon, 2004). A possible alternative interpretation of these, and related, findings is discussed later in this paper.

Bringing the strands together: A more coherent model.

For several years, the working memory hypothesis retained its status as (at least a partial) explanation for children's reading comprehension difficulties, and could be used as an explanation of their problems with inference and integration. But it was important to have a broader overview of the processes that contribute to effective reading comprehension in children, how they differ between good and poor comprehenders, how much variance in reading comprehension they explain, and the extent to which they are dependent on working memory.

More recently, researchers have acknowledged that not only many cognitive skills and knowledge are required for effective reading comprehension, but also many comprehension-specific processes: Not only do words need to be decoded and their meanings retrieved, the meanings of sentences must be derived, information from different sentences integrated, key themes and ideas identified, and inferences made to fill in details and to connect up the text into a coherent representation of the text overall. Furthermore, the reader needs to constantly monitor and reflect on his/her understanding (or lack of) of the text. It is these skills and processes that have been the focus of my later research, for reasons explained in more detail below. In a skilled comprehender these different processes typically go on in parallel, and often without conscious strategic control.

Thus, it is clear that poor text comprehension cannot always be characterised as the result of a decoding problem or some sort of broad cognitive (e.g. memory) problem. There appears to be a separate comprehension component to reading. The relation between language comprehension and word reading has been quantified in The Simple View of Reading (Gough & Tunmer, 1986), which argues that reading comprehension ability is the product of word reading and listening comprehension. From this model, four types of reader can be predicted, including the difference poor readers, identified by Cromer.

A helpful way of thinking about these different aspects of reading, and the processes that contribute to them, is the Rope Model: Scarborough (1981) (the basic Rope Model is shown in Figure 2).

FIGURE 2 about here

Skilled reading can be conceptualized as the weaving together of the two main strands – efficient word reading and language comprehension – and the focus from here on will be on the threads of skills that contribute to the language comprehension strand. The text representation will be very similar whether it is formed via reading or listening though, of course, reading comprehension requires the additional skill of being able to recognize the words.

Models of text comprehension

The language comprehension strand of reading was initially described without much consideration of the skills and processes that might contribute to it. Although, since the beginning of comprehension research, people have agreed that the end product of comprehension is an integrated mental model. But it is only relatively recently that researchers have tried to specify the processes that result in that model. In other areas of reading, such as word decoding and syntactic processing, there are distinct, opposing and empirically falsifiable theories as to how words are read or sentences are understood. In contrast, distinct theories of reading comprehension are more elusive. They might be better described as "frameworks" (see Perfetti & Stafura, 2017). This situation most likely arises because text comprehension is much more complex than other aspects of reading, such as word recognition and sentence parsing, because it requires the successful orchestration of a number of different cognitive skills and abilities, a viewpoint that was only beginning to emerge in the late 1970s (mostly inspired by Bransford's work, as mentioned earlier).

Furthermore, comprehension is not all-or-nothing, in the sense that it either happens or it does not: a reader can have an incomplete, or slightly wrong, interpretation of a text, or one that differs slightly from that of another reader. Thus, although we can state that a person can or cannot accurately read certain words in a text, it does not make sense to say that a certain person makes inferences from text or does not; the criteria for adequate text comprehension are much more subtle. Somewhat different questions are appropriate in relation to text comprehension. So, rather than trying to assess whether or not a particular reader makes inferences, we can ask under what circumstances (text variables, person variables) that reader makes certain types of inference, and whether they make sufficient and appropriate inferences to ensure that they comprehend the text adequately. We can also ask when those inferences are made: immediately as the text is read, or later on when prompted by a question or post-reading task.

Although there are many frameworks or models of skilled text comprehension (see McNamara & Magliano, 2009, for a review), none focus specifically on comprehension development (see Cain & Barnes, 2017, for further discussion on this point). The lack of a single comprehensive theory that can adequately account for text comprehension, its development, and individual differences therein, suggests that it is more productive to think of reading comprehension in terms of the contributory processes, their relative contributions and inter-independencies, and which of these might be lacking in children who experience comprehension problems. In the following sections, I will present more recent research findings, which help to fill in the details of the language strand of the Rope Model.

Processes of comprehension

These processes beyond the word and sentence level are often termed “higher-level” processes, as opposed to the “lower-level” processes such as word reading and access to word meanings. In the studies outlined above, these word-level factors were controlled for (typically by matching groups). However, this matching will have excluded children who had poor comprehension and poor vocabulary, and so perhaps an important explanatory variable (vocabulary) was being ignored. This seems not to be the case: Even though vocabulary (typically breadth of vocabulary – see below) and reading comprehension are typically highly correlated, and there is a causal link between vocabulary skills and later reading comprehension (Quinn et al., 2014; Verhoeven & Leeuwe, 2008), problems with reading comprehension can arise even in the presence of a good vocabulary. Furthermore, variation in comprehension ability cannot be accounted for in terms of variation in vocabulary alone (Oakhill & Cain, 2012). Such findings led to the hypothesis that other, higher-order text-level processes must be important. Subsequent studies focused on three main areas: inference and integration, comprehension monitoring and text structure understanding.

Integration and inference. As mentioned above, some earlier studies demonstrated that poor comprehenders had difficulties with the integration of information in text, and this is one area that has been followed up more recently and in more detail. In addition to the integration of information across sentences (resolution of anaphoric links, use of causal and temporal connectives, etc.), the information in the text must be integrated with prior knowledge, and this integration will support inference making. There is now evidence that poor comprehenders have problems with both local cohesion and global coherence inferences. For instance, in Cain and Oakhill’s (1999) study, the children read a number of short stories each followed by several questions. Some questions tapped their literal comprehension, and there were two types of inference question. The poor comprehenders performed poorly on both local cohesion and global coherence inference questions, but did not have any problems answering the literal questions. More recent research has demonstrated that, even once word decoding skills and vocabulary are taken into account, higher-level skills such as integration and inference making contribute to children's reading comprehension, both within and across time (Bowyer-Crane & Snowling, 2005; Oakhill, Cain & Bryant, 2003; Oakhill & Cain, 2012). Children with reading comprehension problems do not generate as many necessary local cohesion and global coherence inferences as their peers (Cain & Oakhill, 1999; Cain, et al., 2001; Oakhill, 1984). So, for instance, given the following excerpt from a short story:

> On their way back, the weather changed. They ran all the way, but were soaked through when they got home. First, they changed out of their wet things. Then they put the clothes in the washing basket. Then they put on some clean, dry clothes and made some hot chocolate. The warm drink soon made them feel better.

A local cohesion question might be "What drink made Ann and David feel better?" and a global coherence question might be: "What was the weather like on the way home?". The former can be answered by making the anaphoric link between "hot chocolate" and "The warm drink", whereas the latter can be answered by using cues scattered throughout the text: "soaked through", "wet things", "put on … dry clothes".

Three main factors seem to be particularly important in supporting inference making: Knowledge, the reader's standard for coherence, and working memory. The role of these factors in supporting inferences will now be addressed.

Of course, vocabulary knowledge and more general background knowledge are crucial to support inference making, as even very simple inferences cannot be made if the reader does not have the requisite background knowledge. But, even when good and poor comprehenders are matched for background knowledge, the poor comprehenders still show specific difficulties with inference questions (Cain, Oakhill, Barnes & Bryant, 2001), which can probably be attributed to failures to activate and use relevant knowledge. Further evidence that supports a similar conclusion comes from a study by Elbro and Buch-Iversen (2013). They showed that training with graphic organisers, which focused on the contribution of background knowledge for text comprehension, improved children's ability to make global coherence inferences (i.e. those inferences that require the integration of background knowledge with information in the text).

Variation in children's standard for coherence (van den Broek, Risden, & Husebye--Hartman, 1995) may also help to explain inference making differences between good and poor comprehenders. Standards of coherence can be described as the criteria for coherence that a reader applies while reading a text. These standards are criteria that a reader can adopt (either explicitly or implicitly) to suit a particular reading task, depending on their required level of understanding. So, for example, a reader who is trying to find a particular fact in a text might skim the text for specific relevant words, and have a very low standard of coherence for the text overall as his/her aim is not to read to understand the text in general. Children who are good comprehenders are sensitive to different task goals. For instance, when told that they will be tested on their memory for the content of a text, they take longer to read it, and also remember more of its content compared to a 'reading for pleasure' goal. In contrast, poor comprehenders do not adjust their reading in response to different goals (Cain, 1999). This ability to set appropriate standards for coherence may we be related to children's ability to monitor their own comprehension (see below).

Finally, as outlined above, relations between children's inference skills and working memory have been shown (Barnes, Faulkner, Wilkinson, & Dennis, 2004; Cain, Oakhill, & Bryant, 2004; Cain, Oakhill, & Lemmon, 2004).

Monitoring. A second main area of investigation was the relation between comprehension monitoring and reading comprehension. Typical comprehension monitoring tasks (error-detection tasks), and the results of group comparisons on such tasks, have already been described briefly above. For example, a study by Oakhill, Hartt and Samols (2005) showed that children identified as poor comprehenders have difficulties in detecting internal inconsistencies in texts (of the sort described above), but have particular difficulties when the inconsistencies are not in adjacent sentences in the text (i.e. the information that had to be integrated in order for the inconsistency to become apparent was separated by several sentences in the texts). In this study, an assessment of numerical working memory was also included, to explore the relation between comprehension monitoring and performance on the error detection task. However, the relation between the children's performance on the working memory tasks and performance on the comprehension monitoring task (a task that requires both integration and anomaly detection), was not as expected. It was predicted that working memory would be a stronger prediction of error detection in the far than in the near condition. In fact, working memory was the stronger predictor of error detection performance in the near condition and comprehension skill in the far condition. A possible explanation of this finding is that the poor comprehenders do not set up an adequate text representation (or mental model) as they read, so that information later in the text is not necessarily recognized as being in conflict with information presented earlier because the representation of the earlier text was inadequate or incomplete. Thus, the results are consistent with the idea that working memory supports local text integration, but that integration across wider stretches of text is more dependent on the quality of the reader's mental model thus far.

Structure awareness. The third main area of investigation was understanding of text structures (particularly story structures). A range of tasks have been used to explore children's understanding of text structures. This range is noteworthy because some of the tasks did not require reading (understanding and re-telling a story from a pictures sequence;

telling a story based on a given title). We have found that poor comprehenders perform worse on a task that requires them to reorganize a set of jumbled sentences into a coherent story ("story anagram task") (Oakhill, Cain & Bryant, 2003). They are less good at picking out the main point of a picture story from a set of alternatives (Yuill and Oakhill, 1991). Furthermore, in an oral story production task (where they were given a topic on which to base their story), poor comprehenders produce less well-structured stories - their stories are have poorer global coherence, and often lack a main point (Cain & Oakhill, 1996). More generally, the quality of children's story structures is linked with their reading comprehension and reading difficulties (Cain, 2003; Shapiro & Hudson, 1997). Cain (2003) found that 7- to 8-year-old children with reading comprehension difficulties were poorer at telling well-structured stories than their peers.

The link between story structure awareness and reading comprehension is further supported by other findings. For example, poor comprehenders have been found to be less likely than their peers to produce continuations of stories that fit in with the structure of the stories (Englert & Thomas, 1987). Even with informational text, poor structure awareness is linked to poorly structured understanding and recall even with well-structured texts (Taylor & Samuels, 1983).

Causal influences on comprehension development

By the end of the 1990s, there was substantial evidence that certain skills were correlated with reading comprehension, and more information about the potential strands in the Rope Model, but little or no evidence for causal links. Fundamentally, we still had only a catalogue of processes on which poor comprehenders and skilled comprehenders differ, and no evidence as to whether or not any of these processes were causally implicated in reading comprehension. It is important to know about potential causal links because if the point is to help poor comprehenders to improve their comprehension (and young children to avoid comprehension problems) then there is a need to identify, train, and support the skills that have a causal influence on reading comprehension. It is well known that lots of reading helps develop many skills and abilities, including reading comprehension (e.g. Guthrie & Humenick, 2004) so there is little point in investing time and resources teaching children skills that they would develop simply by being encouraged to read more. Of course, it is most unlikely that cause flows in one direction only: it is much more likely that there is reciprocity between reading comprehension and the component skills and processes. But even if there is bi-directionality, teaching children to use skills and processes is likely to improve their reading comprehension over and above just reading.

We took a three-pronged approach to exploring causality, using comprehension-age-match studies, training studies and a longitudinal study. Research using comprehension-age match and longitudinal designs, and training studies provide evidence that difficulties with inference are causally implicated in children's reading comprehension, and particularly reading comprehension breakdown. Comprehension-age-match studies rule out reverse causality, but cannot prove causality (a detailed discussion the interpretation of this design can be found in Cain, Oakhill, & Bryant, 2000). Poor comprehenders make fewer inferences than younger children matched for absolute level of comprehension skill on a standardized reading test (a comprehension-age match group design), which suggests a causal relation between inferences and comprehension skill, rather than vice-versa (Cain & Oakhill, 1999) and similar results have been found for understanding of text structure: Cain (2003) found that 7- to 8-year-old children with reading comprehension difficulties were even poorer at retelling well-structure stories than 6- to 7-year-old children who were matched on level of reading comprehension.

The second type of study used to address causality is a training study. The assumption of such designs is that, if less-skilled children are trained on a target skill, and their reading comprehension subsequently improves (in comparison to an appropriate control group), then it is very likely that that skill was causally implicated in their improvement in comprehension. Training poor comprehenders to be more aware of the need to make inferences results in gains on standardised assessments of reading comprehension, in addition to better inference skills (Yuill & Oakhill, 1988). A different approach uses graphic organisers to make students aware of their own contributions to inferences, by drawing on their background knowledge (Elbro & Buch-Iversen, 2013). In both cases, increased performance on a standardised comprehension assessment was demonstrated. Training of text structures has also been shown to improve not only understanding of text structures, but also reading comprehension (e.g., Hebert et al. 2016).

Finally, we have evidence that all three sets of skills are causally implicated in comprehension development from a longitudinal study: inference skills, story structure understanding and comprehension monitoring predict reading comprehension longitudinally even when the autoregressive effect of reading comprehension, verbal IQ, vocabulary, and other potentially contributory variables have been controlled (Oakhill & Cain, 2012). Thus, the contribution of these skills is not simply because of the early association with comprehension, providing good evidence that they are causal factors (see de Jong & van der Leij, 2002, for support for this argument).

In our longitudinal study, a supplementary finding (which is highly consistent with the Rope Model) was an almost complete dissociation between predictors of word reading and predictors of comprehension (see also Muter, Hulme, & Snowling, 2004). This pattern suggests that comprehension does not necessarily develop automatically once word decoding is proficient (in line with my early findings), but that it is dependent on different skills, and may need specific teaching. Indeed, our early work (Yuill & Oakhill, 1988) has provided the basis for a very successful inference training programme in the UK (*Inference Training:*

Whatmuff, 2013), and many of our studies are cited as the inspiration for comprehension training programmes in the UK (e.g., *the Reading for Meaning Project*: Clarke et al., 2010) and elsewhere (e.g. *LEE Comprehensivamente*, Fonseca et al., 2011, in Argentina).

It should be noted that this longitudinal study was carried out with a population who showed a normal distribution of word reading and reading comprehension – not the selected groups that we had used previously. We have moved towards using regression-based methods rather than comparing groups, because the group comparison ignores the fact that there is a real continuum of skills.

In the longitudinal study, verbal working memory assessments were found to be significantly correlated with reading comprehension (both within each time point, and across time). However, it is noteworthy that, when pitted against the other variables, verbal working memory did not predict significant variance in reading comprehension (either within, or across, time), thus providing no evidence for a causal effect. Perhaps, as suggested by those earlier results (that working was more important for close error detection and comprehension important for more distant error detection), working memory is important for the integration of local information, but what is important for understanding the text as a whole is that the reader starts to build a mental model right from the outset so that new information can be fitted into that model. After all, even a short text is too long to be held in working memory as traditionally thought of (though see Ericsson & Kintsch's idea of "long-term working memory": 1995).

Based on the above findings, we can now elaborate on the Rope Model to represent the different skills and processes that contribute to reading comprehension. Figure 3 shows an adaptation of Scarborough's original Rope Model to reflect this more detailed thinking about the components of reading comprehension.

FIGURE 3 about here

The return of vocabulary: A deeper look

A further issue is that the relation between vocabulary and comprehension might depend on how vocabulary is assessed: More recently, we have considered the role of different aspects of vocabulary knowledge and their relation to reading comprehension. The most commonly used assessments of children's vocabulary measure vocabulary knowledge at relatively shallow levels. For example, in the British Picture Vocabulary Scale (BPVS: MacGinite et al., 2000) the task is simply to match a spoken word with a picture (from a choice of four). Such tests are considered to measure *breadth of vocabulary*: Roughly, how many words are known. But more recently, authors such as Ouellette (2006) and Tanenbaum, Torgesen and Wagner (2006) have also considered readers' *depth of vocabulary*: what they know about the words in their vocabulary. Depth would be assessed by measuring understanding of multiple meanings, or the ability to provide word definitions. These distinctions can be related to Perfetti's *Lexical Quality Hypothesis,* which emphasizes the quality (not just quantity) of lexical representations. Perfetti argued that a low-quality code, which is retrieved with effort, can interfere with comprehension processes. Lexical Quality includes a range of knowledge about word forms (phonology, orthography, grammar) and also meaning. The focus here is on that last aspect: the quality of a word's meaning representation. Perfetti and colleagues (e.g. Perfetti, 2007; Perfetti, Yang & Schmalhofer, 2008) have demonstrated that the availability of associative links between words and concepts – the consequence of a rich (deep) vocabulary – might aid comprehension by supporting inference making in adults.

Recent research with children suggests that measures of vocabulary at greater depth are predictors of reading comprehension over and above measures of vocabulary breadth (Ouellette, 2006; Tanenbaum et al., 2006). Depth of vocabulary knowledge is likely to be more important than breadth in supporting inference and integration in particular, because rich and well-connected semantic representations of words will permit the rapid activation not only of a word's meaning but also the meanings of related concepts. It is plausible that, in children, a rich (deep) vocabulary knowledge will support inference making in comprehension because many of the local cohesion and global coherence inferences in text are dependent on semantic links between words in the text. This activation of a semantic network can then provide the basis for many of the inferences that are crucial for the construction of a coherent representation of a text: a sort of "scaffold" for the mental model.

Evidence that vocabulary supports inference making in children.

A recent study with children provides evidence that depth of vocabulary knowledge is strongly related to making global coherence inferences from text. We showed that depth, but not breadth, of vocabulary knowledge was an important predictor of global coherence inferences, and that this relation held even when word reading skill and literal memory for the text had been taken into account (Cain & Oakhill, 2014). A further study of children aged 6 to 10 years also indicates that vocabulary is more predictive of ability to make global coherence inferences than inferences that link adjacent sentences in text (Currie & Cain, 2015).

We also have evidence from a rather different source that good and poor comprehenders differ in their ability to make semantic associations, as shown by their propensity to derive themes from word lists. Weekes, Hamilton, Oakhill and Holliday (2008) used the DRM (Deese, 1959; Roediger & McDermott, 1995) false memory paradigm with groups of skilled and less-skilled comprehenders. In this paradigm, the children were required to listen to, and to try to memorise, a short list of words. An example word list was: *rest, bed, snooze, dream, tired, blanket*. They then completed a recognition test in which they were asked to differentiate between words that had/had not occurred in the lists they had been read. The good comprehenders were more likely than poor comprehenders to falsely claim that sleep had been in the original list in the example given above, even though they did not have poorer word memory more generally. This result can be taken as an indication that the good comprehenders are more likely to automatically derive "themes" from the word lists (even though this is not a requirement of the task) and, thus, falsely remember words that capture the theme of the list.

This propensity to derive themes from word lists might well support inference making in text comprehension because very often the main theme or the setting of a text can be derived from a number of specific words. For example, if you were to read a text that contained the words: *trolley, shelves, tins, packets, aisle, scan, bags, pack, till*, you might reasonably infer that the text is situated in a shop or supermarket. No one of those words in isolation will support that inference about the setting, but taken together they connect up to provide a coherent overall indication of it.

The results we have reported suggest that the child's depth of vocabulary knowledge might underpin the ability to make global coherence inferences from a text, in particular. The way in which this might work can be illustrated by considering the following snippet of text:

> Today was the big match, the last game of the season. There was only a minute of the game left and neither team had scored yet. The crowd watched in silence as Jake took the penalty shot. The goalkeeper missed the ball and the crowd cheered and roared. Jake had scored the winning goal. His team mates were very happy.
>
> After the match, both teams went back to the changing room. Jake got his shampoo and towel and had a shower. Afterwards, he put his cleaning stuff away in his locker and went home.

In order to answer that question, "What sport was Jake playing?", the reader needs to appreciate that several words taken together (*match, game, scored, penalty shot, goalkeeper, team* and *winning goal*) indicate that it is football. If the reader can make associations between the meanings of different words in the text quickly and easily, he or she will very easily be able to infer what is going on in the story, and the question will be trivially easy.

Speed of meaning access and inference making.

The idea that the automaticity with which meanings can be accessed is important in reading comprehension dates back a long way: Early models of reading (e.g., LaBerge & Samuels, 1974) emphasized the importance of fluency and automaticity of access to word meanings. As mentioned previously, one of my early studies demonstrated a relation between comprehension skill and semantic access (as assessed by naming speed for pictures). More recent correlational studies have shown that good and poor comprehenders differ on measures of semantic fluency, e.g. ability to rapidly produce a number of instances of a category (Nation & Snowling, 1998). Furthermore, Oakhill, Cain, McCarthy and Field (2012) found a strong and specific link between speed of semantic access in vocabulary demanding tasks (synonym and hypernym[5] judgment tasks) and reading comprehension (measured by a standardized task). The link was specific because it was not entirely mediated by word reading skill, or by knowledge about words (assessed by a synonym and hypernym production task), and neither was it related to a simple association between comprehension skill and generally faster response times in a control task that required non-semantic identity judgments (i.e. the ability to say whether two words were orthographically identical or not).

[5] Hypernyms are terms for superordinate categories. For example, rain/weather, dog/mammal, chair/furniture.

These findings indicate that not only must readers understand the meanings of words to comprehend a text, but that knowledge of the network of meaning relations between words, and rapid access to the semantic representations of words, are both also important contributory factors. Since comprehension happens in real time, appropriate meanings and associations of words need to be accessed very rapidly; otherwise the reader will have moved on in the text and the opportunity for semantic information to support inference and integration will have been missed.

It is not necessarily the case, however, that all children's inference making is rapid and automatic. Indeed, other work (Cain & Oakhill, 1999) has shown that poor comprehenders are able to make appropriate inferences strategically, in response to questions, if they are directed back to the text and, in particular, if the parts of the text that support the inference are pointed out to them. Such strategic inferences might also be a product of children's effective comprehension monitoring, which alerts them to the fact that an inference is required.

Concluding comments

Despite decades of research into children's reading comprehension and difficulties, there are still numerous unknowns. We are only beginning to explore how children's representations at the word level help support their text integration and inference skills, and a more detailed exploration of different aspects of vocabulary knowledge and their contribution to reading comprehension would be fruitful. Furthermore, the components that contribute to effective comprehension might be further decomposed as we have tried to do with inference (which is dependent on vocabulary, particularly depth of vocabulary, knowledge of the text domain, working memory, metacognition, and drive for coherence). Whether or not improvements in inference skill, and reading comprehension more generally, could best be improved by teaching inference skills directly, or by supporting some of the 'finer threads' that underlie inference skill is an open question.

It is only in the last decade or so that programmes have begun to emerge that aim to support the development of reading comprehension from a young age. But it is still too often the case that teachers do not realise that reading comprehension will (often) need to be taught alongside word decoding skills, or that they believe that it is something that should wait until children are efficient at decoding words. The importance of developing children's language skills from a very young age (see, e.g., Beck & McKeown, 2010) to support their subsequent reading comprehension cannot be overemphasised (the teaching of reading need not always entail reading!).

Of course, children with word reading problems might also have comprehension problems (Cromer's deficit poor readers) and this is a neglected group, particularly since their comprehension difficulties might not become apparent for some years because their word-reading difficulties are the most obvious problem. Thus, it is important to consider the comprehension ability and needs of all children, including those whose word recognition is delayed, and to provide support when needed. Early screening using tasks that assess listening comprehension, or understanding of cartoon sequences, could provide an indication as to which children are likely to go on to have problems with reading comprehension later, regardless of their word reading skills (see, e.g. Kendeou, van den Broek, White, & Lynch, 2009).

Research into children's reading comprehension – both development and difficulties – has come a long way in the last few decades, and we now have a relatively detailed understanding of the skills and processes that contribute to it. Notwithstanding these developments, there is more to be learned about the interrelations between different skills and processes and, in particular, more to be done to find the best ways to implement the research findings in practice.

Acknowledgments

This article is based on my Distinguished Scientific Contribution Award Address presented at the 29th annual meeting of the Society for Text and Discourse, in New York, 2019. In both the talk and subsequent paper, my aim has been to present an overview of some of my research findings along with milestones along my research path. In doing so, I hope to not only summarise the development of, and findings from, my own and related studies, but also to convey the message that the research path is not always straight and smooth: studies often don't turn out as expected, and potentially important results might be overlooked. I would particularly like to acknowledge funding from the UK Economic and Social Sciences Research Council (ESRC), the EU, the British Academy and the British Council. I thank my main long-term collaborators for their inspiration and support (in alphabetical order): Peter Bryant, Kate Cain, Carsten Elbro, Alan Garnham, Phil Johnson-Laird and Nicola Yuill, and thank Kate Cain and Carsten Elbro for providing comments on an earlier draft of this paper.

References

Anderson R. C., Pichert, W., Goetz., F. T., Shallert, D., Stevens, K. V. & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15*, 667-79.

Barnes, M. A., Faulkner, H., Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language, 89*, 47-56.

Beck, I. L., McKeown, M. G. (2001) Text Talk: Capturing the Benefits of Read-Aloud Experiences for Young Children. *The Reading Teacher, 55(1),*

Bishop, D. V. M. (2003). *Test for the reception of grammar Version 2.* Pearson International.

Bowyer-Crane, C. & Snowling, M.J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology, 75*, 189–201.

Bransford, J. & McCarrell, N.S. (1975). A sketch of a cognitive approach to comprehension: Some thoughts about what it means to comprehend. In W.B. Weimar & D.S. Palermo (Eds.). *Cognition and the Symbolic Processes*, pp. 189-229: Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cain, K. (1999). Ways of reading: How knowledge and use of strategies are related to reading comprehension. *British Journal of Developmental Psychology, 17*(2), 293–309.

Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology, 21*(3), 335–351.

Cain, K. & Barnes, M.A. (2017). Reading Comprehension: What develops and when? In K. Cain, D.L. Compton & Parrila, R.K. (Eds.), *Theories of Reading Development.* John Benjamins. https://doi.org/10.1075/swll.15.15cai

Cain, K. and Oakhill, J.V. (1996) The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology, 14,* 187-201.

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure. *Reading and Writing: An Interdisciplinary Journal, 11*(5–6), 489–503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability and their relation to knowledge. *Memory and Cognition, 29*(6), 850–859.

Cain, K., Oakhill, J. V., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 671–681.

Cain, K., Oakhill, J. V., & 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*(4), 671–681.

Clarke, P.J., Snowling, M.J., Truelove, E. and Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: A randomised controlled trial, *Psychological Science.* 21, 8, 1106-16.

Cromer, W. (1970). The difference model: a new explanation for some reading difficulties. *Journal of Educational Psychology, 61*, 471-483.

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*(1), 51–77.

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. http://dx.doi.org/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.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9,* 561-584.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22.

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading, 17*(6), 435–452.

Englert, C. S., & Thomas, C. C. (1987). Sensitivity to text structure in reading and writing: A comparison between learning disabled and non-learning disabled students. *Learning Disability Quarterly, 10*(2), 93–105.

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

Fonseca, L., et al., (2011) *Programa LEE Comprensivamente.* Buenos Aires: PAIDOS

Golinkoff, R. M. & Rosinski, R. R. (1976). Decoding, Semantic Processing, and Reading Comprehension Skill. *Child Development, Vol. 47,* 1, 252-258.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education, 7(1),* 6-10. doi:10.1177/074193258600700104

Guthrie, J. & Humenick, N. (2004) Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle & V.

Chhabra (Eds.), *The Voice of Evidence in Reading Research,* pp. 329-54. Baltimore: Brookes Publishing.

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology*, 108(5), 609-629. doi:10.1037/edu0000082

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*, 127-160. doi:10.1007/BF00401799

Johnson-Laird, P.N. & Wason, P. (1977). *Thinking: Readings in Cognitive Science*. Oxford University Press.

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*(4), 765-778.

Laberge, D. and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests (4th ed.).* Itasca, IL: Riverside.

McNamara, D. S., & Magliano, J. (2009). Toward a Comprehensive Model of Comprehension. *Psychology of Learning and Motivation: Advances in Research and Theory, 51, 297-384.* DOI:10.1016/S0079-7421(09)51009-2

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665–681.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word--recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*(1), 85–101.

Neale, M. D. (1997). *The Neale analysis of reading ability–Revised (NARA–II)*. Windsor: NFER-Nelson.

Oakan, R., Wiener, M., & Cromer, W. (1971). Identification, organization, and reading comprehension for good and poor readers. *Journal of Educational Psychology, 62*(1), 71-78. http://dx.doi.org/10.1037/h0030778

Oakhill, J. V. (1982). Constructive processes in skilled and less-skilled comprehenders' memory for sentences. *British Journal of Psychology, 73*(1), 13–20.

Oakhill, J.V. (1984) Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54*, 31-39.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91–121.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*(4), 443–468.

Oakhill, J. V., Cain, K., McCarthy, D., & Nightingale, Z. (2012). Making the link between vocabulary knowledge and comprehension skill. In A. Britt, S. Goldman, & J.-F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. 101–114). Hoboken, NJ: Routledge.

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing, 18*, 657–686.

Oakhill, J.V., Yuill, N.M. and Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, 9*, 80-91.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554–566.

Paris, S.G. & Carter A.Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology, 9,* 109-113.

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

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

Perfetti, C. A., Finger, E., & Hogaboam, T. W. (1978). Sources of vocalization latency differences between skilled and less skilled young readers. *Journal of Educational Psychology, 70*(5), 730–739. https://doi.org/10.1037/0022-0663.70.5.730

Perfetti, C., & Stafura, J. Z. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies in Reading, 18,* 22-37.

Perfetti, C. A., 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

Proctor, C. P., Silverman, Shapiro, B. K., & Hudson, J. A. (1997). Coherence and cohesion in children's stories. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 23–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, 46, 247-253. doi:10.2307/1128859

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy* (pp. 97-110). New York: Guilford Press.

Singer, H. & Ruddell, R. (Eds.) (1976). *Theoretical models and processes of reading* (pp. 259-271). Newark, DE: International Reading Association.

Steiner, R., Wiener, M., & Cromer, W. (1971). Comprehension training and identification for poor and good readers. *Journal of Educational Psychology, 62*(6), 506-513. http://dx.doi.org/10.1037/h0032035

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*(4), 381–398.

Taylor, B. M., & Samuels, S. J. (1983). Children's use of text structure in the recall of expository material. *American Educational Research Journal, 20*(4), 517–528.

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

Weekes, B. S., Hamilton, S., Oakhill, J. V., & Holliday, R. E. (2008). False recollection in children with reading comprehension difficulties. *Cognition, 106*(1), 222–233.

Yuill, N. M., & Oakhill, J. V. (1988). Effects of inference awareness training on poor reading comprehension. *Applied Cognitive Psychology, 2*(1), 33–45.

Yuill, N. M., & Oakhill, J. V. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge: Cambridge University Press.

Yuill, N. M., Oakhill, J. V., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80*(3), 351–361.

Weekes, B. S., Hamilton, S., Oakhill, J. V., & Holliday, R. E. (2008). False recollection in children with reading comprehension difficulties. *Cognition, 106*(1), 222–233.

Whatmuff, A. (2013, revised 2016) *Inference Training.* Leicester City Council: Primary Strategy Team. LEE.