Explaining Reading Comprehension in Children with Developmental Language Disorder: The Importance of Elaborative Inferencing.

Sheila M. Gough (corresponding author)

Department of Psychology, University of Roehampton, London, United Kingdom.

Olympia Palikara

School of Education, University of Roehampton, London, United Kingdom.

Rebecca M. Lucas

Department of Psychology, University of Roehampton, London, United Kingdom.

Abstract

Purpose: Reading comprehension is a key indicator of academic and psychosocial outcomes. Children with Developmental Language Disorder (DLD) tend to find reading comprehension challenging. This study aimed to explore the literal and inferential (cohesive, elaborative and lexical) comprehension of children with DLD, their typically developing (TD) peers and, uniquely, a group of Low-Language (LL) children. ***Method:*** Children aged 10-11 with either typical development (n = 16), LL (n = 14) or DLD (n = 14) were recruited from eight primary schools. They completed a battery of standardized language and literacy assessments. Responses to literal and inferential questions on the Wechsler Individual Achievement Test-UK (WIAT-UK; Wechsler, 2005) were analyzed. ***Results:*** A disproportionate difficulty in answering inferential relative to literal questions was found for the DLD group compared to LL and TD peers. Children with DLD were significantly poorer at elaborative inferencing than both their LL and TD peers, but there were no group differences in cohesive or lexical inferencing. There was a significant positive association between inferencing ability and vocabulary knowledge, single word reading accuracy, grammatical skill and verbal working memory. The importance of single word reading accuracy was especially evident as a partial mediator of the relationship between vocabulary knowledge and inferencing ability. ***Conclusions:*** These results indicate that interventions targeting the reading comprehension of children with DLD should focus on elaborative inferencing skill. There are also clinical implications as the development of new standardized assessments differentiating between inference types is called for.

Key words: Developmental Language Disorder, low language, reading comprehension, elaborative inferencing, inference deficit, inferencing

Explaining Reading Comprehension in Children with Developmental Language Disorder: The Importance of Elaborative Inferencing.

Developmental Language Disorder (DLD), previously known as Specific Language Impairment (SLI)[1], is a neurodevelopmental disorder which affects approximately 7.5% of children (Norbury, Gooch, Wray, Baird, Charman, Simonoff, Vamvakas & Pickles, 2016; Tomblin, Records, Buckwalter, Zhang, Smith & O'Brien, 1997). Language impairments are evident across language areas (e.g. phonology, semantics and syntax) and modalities (i.e. spoken and written) and these difficulties can be receptive, expressive or mixed (American Psychological Association, 2013; Bishop, Snowling, Thompson & Greenhalgh, 2016). Children and adolescents with DLD tend to have poorer academic attainment and psychosocial well-being than their typically developing (TD) peers (Conti-Ramsden, Bishop, Clark, Norbury & Snowling, 2014; Dockrell, Lindsay, Palikara & Cullen, 2007) and their needs are pervasive. A key predictor of outcomes is reading competence, especially reading comprehension (Conti-Ramsden, Durkin, Toseeb, Botting & Pickles, 2017; Cromley, 2009; Hernandez, 2012; Vilenius-Tuohimaa, Aunola & Nurmi, 2008). Given the importance of reading comprehension for optimal academic and psychosocial outcomes, it is imperative that we increase our understanding of factors associated with reading comprehension for children with DLD. One aspect of reading comprehension that we have limited knowledge of is inferencing; while we know that children with DLD find inferencing more challenging than their typically developing peers (Lucas & Norbury, 2015), we know little about their experience with different types of inferencing. This study aimed to explore cohesive, elaborative and lexical inferential comprehension and literal comprehension in a sample of

[1] Practitioners were concerned that a lack of consensus with regards to terminology and criteria was creating a barrier to prevention and intervention services for children with language disorder. A multi-disciplinary consortium of experts employed a consensus building model (Bishop, Snowling, Thompson & Greenhalgh, 2016). DLD is to be used when the language disorder is not associated with a known aetiology. This was heretofore often referred to as 'SLI'.

children with DLD, children with Low-Language (LL) proficiency and their TD peers. The knowledge generated can feed into the development of evidence based targeted interventions to improve reading comprehension, which are currently limited (Brooks, 2016).

Models of Reading Comprehension

Reading is a highly complex skill, but it comprises two core components: decoding and comprehension (Hoover & Gough, 1990). These aspects typically develop in tandem (Gough, Hoover, & Peterson, 1996), but word recognition is critical for successful comprehension. Accordingly, the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) describes reading comprehension as the product of decoding and oral language comprehension, and there is considerable support for this model (Roth, Speece, Cooper & Paz, 1996; Braze, Katz, Magnuson, Mencl, Tabor, Van Dyke, Gong, Johns & Shankweiler, 2016; Compton, Miller, Elleman & Steacy, 2014). Another model of reading comprehension that was developed at a similar time as the Simple View of Reading is Kintsch's (1988) Construction Integration (CI) model. This model defines three sources of input from the text, linguistic input, inference made from linguistic input and general background knowledge. This framework fits well with the three distinct inferencing types identified in this study: cohesive, whereby conclusions are drawn by establishing links between premises within the text, elaborative, whereby conclusions are drawn by adding the background knowledge to information contained within the text and lexical, whereby the meaning of vocabulary is established using the context of the text. From these, the reader gleans understanding, or constructs propositions, and finally integrates propositions into one coherent message. However, neither of these theories accounts for all of the variance in reading comprehension and have been criticized for being too simplistic (Cartwright, Marshall & Wray, 2016; Dixon & Bortolussi, 2013).

These limitations have contributed to the development of the reading comprehension framework (Perfetti, Landi & Oakhill, 2005). This proposes that reading comprehension is a much more complex process, not only underpinned by written word identification and vocabulary, but also by language systems such as syntax, and general knowledge. Thus, reading comprehension impairments can develop as a result of a deficit in any of these domains. For children with DLD, reading comprehension is especially challenging (Bishop, McDonald, Bird, & Hayiou-Thomas, 2009; Palikara, Dockrell, Lindsay, 2011), with approximately 50% having impaired reading comprehension, 15% demonstrating a poor comprehender profile in which reading comprehension is substantially poorer than word recognition, with or without a formal diagnosis (Catts, Fey, Tomblin, & Zhang, 2002; Hulme & Snowling, 2014). Children with LL also find reading comprehension significantly more difficult than their TD peers (Myers & Botting, 2008).

The Role of Inferencing in Reading Comprehension

Reading comprehension involves understanding of explicitly stated information, as well as the ability to make an inference (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Garnham & Oakhill, 1996; Kleeck, 2008). Making an inference requires an individual to go beyond what is explicitly stated and draw a conclusion based on evidence and reasoning. When a skilled reader processes a text, they often use the information within the text and general background information to ‘fill in the gaps’ and achieve greater comprehension. The more skilled the reader, the more inferences they generate (Long, Oppy & Seely, 1997; Prior, Goldina, Shany, Geva & Katzir, 2014). For TD children, there is a positive relationship between inferencing competence and word reading, vocabulary, grammar and working memory, but vocabulary is the critical predictor (Silva & Cain, 2015).

As these aspects are often impaired in children with DLD and LL, it is plausible that inferencing would be especially challenging for these children.

To date there is a paucity of research exploring the ability of children with DLD or LL to make inferences from text, with the sole exception of Lucas and Norbury (2015). The knowledge that we do have of inferencing in children with DLD is largely resultant of research examining inferencing in the oral domain. However, such studies have reported some conflicting findings. Some indicate that children with DLD struggle with both literal and inferential comprehension questions relative to TD peers (Adams, Clarke & Haynes, 2009; Bishop & Adams, 1992; McClintock, Pesco & Martin-Chang, 2014), whilst others report that children with DLD have a selective problem with inferencing (Crais & Chapman, 1987; Dodwell & Bavin, 2008; Karasinski & Weismer, 2010; Ellis Weismer, 1985). In addition, others report that at the group level there are no significant differences in response accuracy on literal or inferential questions between children with DLD and TD children (Norbury & Bishop, 2002). Lucas and Norbury (2015) did, however, examine inferential ability in text comprehension, rather than oral comprehension. They found that children with DLD found inferencing more challenging than their TD peers. This supported findings by McClintock et al. (2014) and Wright and Newhoff (2001); that both TD children and children with DLD were more successful at literal than inferential questions, and TD children performed more accurately than children with DLD on inferential questions in general. No studies have investigated inferencing in children with LL who do not meet the criteria for a clinical diagnosis of DLD.

Inferencing in oral and written domains.

In their examination of inferencing skill, some of the aforementioned studies (Norbury & Bishop, 2002; McClintock et al., 2014) only report group differences, while others have examined predictors of inferencing skill (Botting & Adams, 2005; Adams et al., 2009; Karasinski & Weismer, 2010; Dodwell & Bavin, 2008). Predictors of inferencing skill for children with DLD include vocabulary knowledge (Botting & Adams, 2005), grammatical knowledge (Botting & Adams, 2005), sentence comprehension and age (Adams et al., 2009) and verbal working memory (Karasinski & Weismer, 2010; Dodwell & Bavin, 2008), as well as non-verbal IQ (Botting & Adams, 2005). Lucas and Norbury (2015) found that vocabulary knowledge and verbal working memory were significant predictors of inferencing skill for the sample as whole, but this comprised children with Autism Spectrum Disorder, in addition to those with DLD and TD peers. These studies reported only significant factors, with the exception of Botting and Adams (2005) who reported a teacher/parent completed screen for communication disorder to have a borderline non-significant relationship with inferencing skill.

There are some discordant results from studies investigating inferencing in the oral domain, these include differences found in inferential relative to literal skill and different predictive factors of inferencing skill. Potential reasons for this center around individual differences, participant characteristics, the study materials, and the analyses conducted. Many of the cited studies examine data for the DLD group as a whole without taking into consideration individual differences (e.g. Botting & Adams, 2005; McClintock et al., 2014). Norbury and Bishop (2002) conducted an examination of individual data to determine the percentage of children aged 6-10 who had a disproportionate difficulty with inferential relative to literal reading comprehension questions. They found that 25% of children with DLD had a disproportionate difficulty with inferencing from orally presented stories, compared to only 11% of their TD peers, yet at a group level, the TD and DLD samples did not differ in terms of literal and inferential question response accuracy. Using the same procedure, this time investigating inferencing in text, Lucas and Norbury (2015) found that

58.33% of children with DLD (aged 7-12) found the inferential questions especially challenging, relative to only 12.50% of their TD peers. Other studies (e.g. Botting & Adams, 2005) do not provide details of any comparison made between response accuracy for literal and inferencing questions – this makes it difficult to determine the proportionate difficulty participants had with inferencing skill relative to their skill in answering literal questions.

In examining inferencing skill, studies have used different comparison group criteria. Some studies compare the inferencing skills of children with DLD to age-matched comparisons and also younger TD children matched for expressive language (Dodwell & Bavin, 2008). Other studies matched for receptive narrative (Adams et al., 2009; Ellis Weismer, 1985) or matched with an age group representing the age equivalent language scores of the DLD children (Botting & Adams, 2005; Crais & Chapman, 1987). However further studies matched TD and DLD groups for age alone (Norbury & Bishop, 2002; Wright & Newhoff, 2001). More studies did not match for language ability but for age and sex (McClintock et al., 2014). Different matching criteria changes the relationship between groups and thus makes it difficult to compare findings across studies.

Another challenge to comparing study results is that there are many different measures of reading comprehension, and these vary greatly in terms of the aspects of comprehension they examine (Keenan, Betjemann & Olson, 2008). Thus, there is a lack of consistency in the types of inferences being assessed. There are two main types of inferences: cohesive inferences, whereby conclusions are drawn by establishing links between premises within the text, and elaborative inferences, whereby conclusions are drawn by adding the background knowledge to information contained within the text (Cain, Oakhill, Barnes & Bryant, 2001). It has been established that poor comprehenders are weaker at generating both cohesive and elaborative inferences, relative to their peers capable of skilled comprehending (Cain & Oakhill, 1999). It is uncertain whether this is also the case for children with DLD. Although Botting and Adams (2005) and Norbury and Bishop (2002) distinguished between cohesive and elaborative inferencing (at least in essence if not in terminology), it is not clear from the statistical analyses reported whether response accuracy varies by inferential question sub-type. However, inspection of mean scores indicates that children with DLD may find elaborative inferences more challenging than cohesive inferences (Norbury & Bishop, 2002). Other studies (Adams et al., 2009; Bishop & Adams, 1992; Crais & Chapman, 1987; Dodwell & Bavin, 2008) did not report inferencing ‘type’.

Rationale

Whilst vocabulary predicts inferential competence, successful inferencing could also provide the opportunity for children to cement their knowledge of existing vocabulary. However, to date lexical inferencing has not been the focus of research. Children with poorer language skills (both DLD and LL) find learning new vocabulary inferentially more challenging than their TD peers (Cain, Oakhill & Bryant, 2004; Lucas & Norbury, 2017; Nash & Donaldson, 2005), but it is uncertain whether consolidation or augmentation of existing vocabulary knowledge is also impacted. Children with LL are not as widely present in the literature as children with DLD, but we do have some knowledge of their abilities relative to their TD peers. These children have significantly greater academic difficulties than peers with higher language skills (Myers & Botting, 2008), and yet they do not fit into a diagnostic category. As such, they may not receive the full benefit of the support that a child with a diagnosis would be entitled to, despite being at a similar risk of negative outcomes related to poor reading comprehension ability (Conti-Ramsden et al., 2017). In investigating the links between language and inferencing skill, it is imperative that we not merely look at two polarized groups – children with DLD and their TD peers – but at the full spectrum of needs.

Thus, further research concerning inference generation in children with DLD and children with LL compared to their TD peers is essential. There is a dearth of information on inferencing from text, with an emphasis instead on oral comprehension. Furthermore, as many of the extant studies have not specified the type of inferences assessed nor examined data at the individual level we are not yet able to accurately forecast which children may need most support, nor the optimal form that this should take. This is vital information as improvements in educational support systems can improve outcomes for children with DLD and less obvious language needs (Conti-Ramsden et al., 2017). However, to date educational and psychosocial outcomes for children with DLD are still not as optimistic as for their TD peers; young adults with DLD tend to have lower academic and vocational qualifications than their TD peers and are more likely to be in non-professional occupations (Conti-Ramsden et al., 2017). Reading comprehension is a key predictor of these factors.

Present Study

The current study aimed to explore the literal and inferential comprehension of children with DLD. More specifically it extended previous research by examining cohesive and elaborative inferencing, as well as lexical inferences. It also provided novel data by exploring inference deficits (inference skill relative to literal skill) for each type of inference. Previous research has indicated that competence using linguistic context to resolve lexical ambiguities aligns with language ability (Norbury & Nation, 2011; Norbury, 2005). We therefore predicted that children with DLD would not only be poorer at inferencing (cohesive, elaborative and lexical) than their TD and LL peers, but would also be more likely to have a disproportionate difficulty answering inferential relative to literal questions (cf.

Lucas & Norbury, 2015; Norbury & Bishop, 2002). We also predicted that children with DLD, LL children, and their TD peers would be stronger at generating cohesive inferences than elaborative inferences (cf. Norbury & Bishop, 2002). However, due to the lack of extant research on lexical inferences, we were unable to hypothesize the relative difficulties with this type of inference. Finally, we predicted a complex interaction of predictors: As per the Simple View of Reading, we hypothesised that vocabulary knowledge would predict inferencing skill (cf. Silva & Cain, 2015; Lucas & Norbury, 2015) but that this predictive effect of vocabulary knowledge on inferencing skill would be mediated by single word reading accuracy (cf. Gough et al., 1996). As per the Reading Comprehension Framework, we also predicted that these factors (vocabulary knowledge and single word reading accuracy) would be joined by grammatical skill, and verbal working memory as significant positive predictors of inferencing skill (cf. Lucas & Norbury, 2015; Botting & Adams, 2005; Dodwell & Bavin, 2008; Karasinski & Weismer, 2010).

Method

Participants

65 children (aged 10-11 years) were recruited to the study from year 6 classes in eight primary schools in the south-east of England. The protocol for this study was approved by the Research Ethics Committee at XXX. Verbal assent was obtained from all children and informed, written consent was provided by all parents, teachers, and headteachers.

Children with DLD (n = 14) were currently on their school's special educational need (SEN) register (this is the record of children with special educational needs held by the school; standard procedure in the U.K. school system). They held a label of "Language Disorder" or "Speech, Language and Communication Need", were receiving specialist educational support (e.g. learning support teacher) and their DLD symptomatology was indicated by their teachers through completion of the Children's Communication Checklist 2, (CCC-2; Bishop, 2003b). All groups of participants completed a battery of language assessments to confirm group membership and to assess language skills. These assessments were the 'Recalling Sentences' subtest (measuring expressive and receptive narrative and verbal working memory) and the 'Word Classes' subtest (Receptive and Expressive; measuring vocabulary) of the Clinical Evaluation of Language Fundamentals-IV (UK), (CELF-IV; Semel, Wiig & Secord, 2004), and the Test for Reception of Grammar 2 (TROG-2; Bishop, 2003a) (measuring receptive grammar). All children with DLD obtained a score at or below 1.25SD below the population norm on both a receptive and an expressive language task. These standardized assessments report a score of below 1.25 SD to be indicative of impairment. Peers (n = 51) were recruited from the same schools as the children with DLD. 21 participants with language ability scores more than +1.25SD from the population norm on at least one language measure were excluded from the current study to ensure that the TD group was representative of population norms. The 16 participants who achieved scores within 1.25SD of the population norm on all language tasks and did not have a history of DLD or language delay (according to teacher report) were included as a TD group. A third low language (LL) group (n = 14) included the students who did not have a clinical diagnosis of language disorder but scored at or below 1.25SD on one of the language tasks. Three of these students scored below 1.25SD on the 'Word Classes' receptive subtest of the CELF-IV (Semel et al., 2004); four of these students scored below 1.25SD on the TROG-2 (Bishop, 2003a); and seven of these students scored below 1.25SD on the 'Recalling Sentences' subtest of the CELF-IV (Semel et al., 2004). Thus, they exhibited lower language ability that their peers included in the TD group, but did not score at or below 1.25SD below the population norm on both a receptive and an expressive language task, as per the DLD group.

The DLD, LL and TD groups did not differ in chronological age nor sex (see Table 1). In-line with their group status, the DLD and LL groups had lower scores on the language measures than their TD peers, as well as lower scores on the literacy measures (which are outlined below), see Table 1. Non-verbal cognitive abilities were assessed using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence –Second Edition (WASI-II; Wechsler, 2011), which involved the child selecting a picture to complete a pattern. Similar to other studies, we also found that non-verbal and verbal abilities were highly correlated (cf. Conti-Ramsden, St. Clair, Pickles & Durkin, 2012), such that children with DLD tended to have lower non-verbal ability scores (cf. Dennis, Francis, Cirino, Schachar, Barnes & Fletcher, 2009).

Insert Table 1 here.

Materials and Procedure

Two components of literacy were assessed; single word reading accuracy and passage reading accuracy and comprehension. Single word reading ability was assessed using the sight word efficiency (SWE) and phonemic decoding efficiency (PDE) subtests of the Test of Word Reading Efficiency-2 (TOWRE-2; Torgesen, Wagner & Rashotte, 2011). The administration of the TOWRE involved the child reading two lists aloud, one of real words and one of made-up nonwords. Passage reading accuracy and comprehension were assessed through the Reading Comprehension subtest of the Wechsler Individual Achievement Test-UK (WIAT-UK; Wechsler, 2005). Table 1 reports the standard scores for these reading measures.

For experimental purposes, the administration of the Reading Comprehension subtest of the WIAT-II (Wechsler, 2005) included the ten passages normed for use with 10 and 11 year olds. Therefore all participants began at ‘Toontime Tees’ which is the starting point for children aged 10, and finished at ‘Yukon Gold’, the discontinuation point for children aged 11. This enabled consistency across participants in terms of the comprehension questions administered.

Following the reading of each passage in the WIAT-II (Wechsler, 2005), participants were asked the corresponding comprehension questions. The 34 comprehension questions administered for experimental purposes were analysed by the three authors to identify literal and inferential questions. Questions were categorised as literal if they could be answered by recalling information that was explicitly mentioned in the text. In contrast, if the question could only be answered by the information in the text being used as a basis for reasoning and drawing a conclusion (i.e. the answer had not been directly stated), then it was categorised as inferential. The inferential items were further divided into three types; 1) ‘inferential cohesive', whereby conclusions are drawn by establishing links between premises within the text, 2) 'inferential elaborative', whereby conclusions are drawn by adding background knowledge (life experiences and general knowledge) to information contained within the text, and 3) 'inferential lexical', whereby contextual information is used to reason the definition of key words. This resulted in a total of 18 literal questions and 16 inferential questions, (5 cohesive, 4 elaborative, 7 lexical)[2]. A high degree of inter-rater reliability was found between the three individuals who categorized each question. The average measure Intraclass Correlation (ICC) was .996 with a 95% confidence interval from .994 to .998, $F(38, 76) = 268.21$, $p < .001$.

Each question had a maximum score of 2 points for each correct answer, therefore the maximum total score possible for literal questions was 36, for inferential cohesive it was 10, for inferential elaborative it was 8 and for inferential lexical it was 14. Participants completed the test battery individually over 2 sessions in a quiet room at their school. The Matrix Reasoning subtest of the WASI-II (Wechsler, 2011) was administered in the first session. This was followed by the 'Recalling Sentences' subtest and 'Word Classes' subtest of the CELF-IV (Semel et al., 2004), and then the TROG-2 (Bishop, 2003a). In the second session, the TOWRE-2 (Torgesen et al., 2011) was administered, followed by the Reading Comprehension subtest of the WIAT-II (Wechsler, 2005).

[2] Please contact the authors for access to the detailed categorisation of each WIAT-II (Wechsler, 2011) question into inference type.

Results

Literal and Inferential Reading Comprehension

Due to the unequal number of items for each reading comprehension question type, the raw total scores were transformed into percentages to enable direct comparisons between literal, inferential cohesive, inferential elaborative and inferential lexical questions. All subsequent analysis was performed on percentage accuracy scores.

A 4x3 (Question Type: literal vs. cohesive vs. elaborative vs. lexical Vs; Group: TD vs. LL vs. DLD) mixed ANOVA was conducted, F (3, 123) = 13.05, $p < .001$, $\eta_p^2 = .24$, a small effect size. There was a significant main effect of Question Type, F (1, 41) = 31.77, $p < .001$, $\eta_p^2 = .44$, a medium effect size. Post-hoc analysis indicated that literal questions were answered more accurately than any of the inferential question types (all $p < .001$, please see Figure 1). There was also a significant main effect of Group, F (2, 41) = 16.82, $p < .001$, $\eta_p^2 = .45$, a medium effect size. Posthoc analysis indicated that the DLD group performed significantly lower on all question types than the TD group (all $p < .020$) but no significant differences were found between the TD or DLD and LL groups (all $p > .439$). There was not a significant interaction for Question Type and Group, F (2, 41) = 1.42 $p = .253$, $\eta_p^2 =.07$.

Insert Figure 1 here.

Predictors of Inferencing Competency

To investigate which factors predict inferencing skill for the whole sample, a mediation analysis was conducted exploring the relationship between vocabulary knowledge and single word reading accuracy, as per the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990. This formed the first step of a hierarchical regression analysis. The strong correlation ($r = .96$, $p < .001$) between the expressive and receptive raw scores of the ‘Word Classes’ subtest of the CELF-IV (Semel et al., 2004) justified the creation of a composite score (created by averaging the two raw scores), labelled vocabulary knowledge composite. Likewise, the strong correlation between the TOWRE-2 sight word efficiency (SWE) and phonemic decoding efficiency (PDE) raw scores ($r = .88$, $p < .001$) justified the creation of a single word reading composite (similarly created by averaging the two raw scores). The mediation analysis was performed following the four steps recommended by Baron and Kenny (1986): 1) Establish an effect by showing causal variable to be correlated with the outcome variable, 2) Establish a correlation between causal variable and mediator, 3) Use a regression model to show that when causal variable is controlled, the mediator effects the outcome, 4) Establish full/partial mediation by controlling for the mediator to see if this negates the effect of causal variable on outcome variable. The predictive power of vocabulary knowledge composite on inferencing ability was found to be partially mediated by single word reading composite accuracy, $\beta = .27$, $p = .093$ (please see Figure 2). This partial mediation is demonstrated as the significant effect of vocabulary knowledge composite on inferencing ability becomes non-significant (although not a zero effect) when mediated by single word reading composite accuracy.

Insert Figure 2 here.

Further regression analysis was added to this hierarchical regression. Multiple regression was conducted on all groups, incorporating a wider range of factors as per the Reading Comprehension Framework (Perfetti et al., 2005). In total, four predictor variables were entered into the model: vocabulary knowledge composite; single word reading accuracy composite; grammatical skill (based on raw TROG-2 scores); and verbal working memory (indexed by CELF Recalling Sentences raw scores). The dependent variable was the percentage of inferential questions correctly answered. The total model was significant, $F(4, 43) = 18.92$, $p < .001$, and explained 66% of the variance in the percentage of inferential questions correctly answered. The single word reading composite and verbal working memory were significant predictors of inferencing competence, whilst the vocabulary knowledge composite and grammatical skill did not contribute significant variance (see Table 2).

Insert Table 2 here.

Inferencing Deficits

Figure 1 could indicate that although this sample of children with DLD find inferencing more challenging than their peers, they do not have a disproportionate difficulty with inferencing, relative to their TD and LL peers. However, group means can mask individual differences. To explore this further an ‘inference deficit’ score was created by dividing the percentage of correct inferential answers by the percentage of correct literal answers (cf. Norbury & Bishop, 2002; Lucas & Norbury, 2015). A score of 1 indicates that the child answered inferential questions as accurately as literal questions. Scores 1SD below the TD mean of .85, i.e. scores < .69, are considered to be indicative of a disproportionate difficulty with inferencing relative to the TD peers in this sample. A further ‘cohesive inference deficit’, ‘elaborative inference deficit’ and ‘lexical inference deficit’ were also created to allow a comparison of performance on these factors. The ‘cohesive inference deficit’ was created by dividing the percentage of correct cohesive inferential answers by the percentage of correct literal answers, the ‘elaborative inference deficit’ was created by dividing the percentage of correct elaborative inferential answers by the percentage of correct literal answers and similarly the ‘lexical inference deficit’ was created by dividing the percentage of correct lexical inferential answers by the percentage of correct literal answers (see Table 3). Table 3 includes details on the mean and standard deviation accuracy of each question type by group and also includes details of the number (N) and percentage of participants within each group with an inference deficit in each question type.

Insert Table 3 here.

The percentage of children with an inferencing deficit is displayed in Figure 3. Chi square analyses indicated that there was a marginally significant group difference in overall inference deficit, $\chi 2$ (2, N=44) = 5.65, p = .059, φ = .36. The TD and low-language groups did not differ, $\chi 2$ (1, N=30) = .43, p = .513, φ = .12, , and nor did the DLD and LL groups, χ^2 (1, N=28) = 2.49, p = .115, φ = .30, whilst the DLD group were more likely to have an inferencing deficit than their TD peers, χ^2 (1, N=30) = 5.00, p = .025, φ = .41. However, consideration of inferencing types indicated that this was largely attributable to elaborative inferencing. Children with DLD were more likely to have an elaborative inferencing deficit than both their TD peers, χ^2 (1, N=30) = 10.80, p = .001, φ = .60, and their peers with LL, χ^2 (1, N=28) = 5.14, p = .023, φ = .43 whilst again, the TD and LL groups did not differ, χ^2 (1, N=30) = 1.21, p = .272, φ = .20. There were no group differences in cohesive inferencing, χ^2 (2, N=44) =2.77, p = .251, φ = .25 or lexical inferencing, χ^2 (1, N=44) = 3.81, p = .149, φ = .29.

Insert Figure 3 here.

To investigate which factors predicted the elaborative inferencing deficit, hierarchical regression was again conducted. Vocabulary and single word reading, the two factors relating to the Simple View of Reading, were initially explored. The relationship between the vocabulary knowledge composite and elaborative inferencing deficit was further examined and the predictive power of the vocabulary knowledge composite on elaborative inferencing deficit was found to be partially mediated by the single word reading composite, $\beta = .22$, $p = .217$ (please see Figure 4).

Insert Figure 4 here.

A multiple regression analysis was then added to the hierarchical regression model using the four predictor variables previously described (constituting the Reading Comprehension Framework approach): vocabulary knowledge composite; single word reading accuracy composite; grammatical skill and verbal working memory. The total model was significant, $F(4, 43) = 6.38$, $p < .001$, and explained 40% of the variance in elaborative inference deficit. No individual factor significantly predicted elaborative inferencing deficit, all $p > .05$ (please see Table 4).

Insert Table 4 here.

Discussion

This study investigated the literal and inferential reading comprehension of children with DLD, not only comparing their competency to TD peers, but uniquely, also to children with a 'Low-Language' (LL) profile. Importantly, inferencing was not only considered as a unitary construct, but cohesive, elaborative and lexical inferences were additionally examined separately. The DLD group demonstrated poorer reading comprehension across all four question types relative to their TD and LL peers (who did not differ from one another). For all groups, elaborative inferencing was most challenging, and analysis at the individual level indicated that this was especially the case for children with DLD. An elaborative inferencing deficit was predicted by a model consisting of vocabulary knowledge; single word reading accuracy composite; grammatical skill and verbal working memory. The importance of single word reading accuracy was especially evident as a partial mediator of the relationship between vocabulary knowledge and inferencing ability.

Inferencing in Children with Developmental Language Disorder

The findings support the previous research showing that children with DLD struggle with both literal and inferential comprehension questions (cf. Adams, Clarke & Haynes, 2009; Bishop & Adams, 1992; McClintock et al., 2014). However, they also support the findings that children with DLD experience a disproportionate problem with inferencing relative to TD peers (cf. Crais & Chapman, 1987; Dodwell & Bavin, 2008; Karasinski & Weismer, 2010). The only previous research on inferencing from text within a sample of children with DLD (Lucas & Norbury, 2015), found that children with DLD were more likely than TD children to have a disproportionate difficulty with inferencing. This finding was validated by this study, with comparable percentages of inferencing deficit found: 58.33% of children with DLD and 12.50% of their TD peers in Lucas and Norbury found the inferential comprehension questions more challenging than the literal questions, compared to 66.67% of children with DLD and 18.57% of their TD peers in this study. However, our study further extended this finding by examining cohesive, elaborative and lexical inferencing deficits. Notably, the difference in inferential relative to literal ability for children with DLD was attributable to elaborative inferential questions, rather than cohesive inferential or lexical inferential questions. This study uniquely looked at three different domains: the accuracy of response to literal and inferential questions by group, the predictors of inferencing overall and also the level of 'inferencing deficit' by group so was able to report on each domain.

Predictors of Inferencing Ability

Alone, the vocabulary knowledge composite was found to predict both inferencing ability and the presence of an elaborative inferencing deficit. It is mediated to a large degree by the single word reading composite, which offers support for the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990). Yet in terms of the Reading Comprehension Framework (Perfetti et al., 2005), we expected the vocabulary knowledge composite to explain a larger proportion of the variance outside of the regression model (Silva & Cain, 2015; Botting & Adams, 2005; Karasinski & Weismer, 2010; Dodwell & Bavin, 2008), as we did grammatical knowledge, verbal working memory and the single word reading composite (Botting & Adams, 2005; Adams et al., 2009). The wider range of predictive factors selected for this model was informed by predictors for inferencing skill identified in previous research; vocabulary knowledge composite; single word reading composite; grammatical skill and verbal working memory. This model predicted 66% of the variance in inferencing skill. In both regression models investigating the Simple View of Reading and the Reading Comprehension Framework, the findings could in part be due to the high correlation between some of the variables. Thus while the vocabulary knowledge composite, grammatical knowledge and non-verbal IQ may predict inferencing skill, the significance of this effect after the more significant factors (verbal working memory and the single word reading composite) have been accounted for, is moot. This has been the case in previous research with TD children (Oakhill & Cain, 2012; Silva & Cain, 2015), wherein, receptive grammar skill was found not to be a significant predictor when vocabulary knowledge was taken into account.

This inclusion of a multiple regression analysis of inferencing ability within this research (which moves from the oral domain to look at inferencing within reading comprehension), is very important given the premise that the Simple View of Reading may be too simplistic (Cartwright et al., 2016). It allows the significance of more variables to be recognized, as per the reading comprehension framework (Perfetti et al., 2005). The emergence of verbal working memory and the single word reading composite as significant predictors of an inference deficit implies that a greater array of cognitive processes than posited by the Simple View are involved in reading comprehension. This study measured reading comprehension as per the Simple View of Reading, in that it included measures of decoding and comprehension, but the key finding holds more importance for the reading comprehension framework. Our findings suggest that poor reading comprehension scores for children with DLD is more closely related to elaborative inferencing skills than decoding/comprehension. For this population, it would seem that reading comprehension is underpinned by an ability to draw upon background knowledge (and indeed, to have embedded life experience into background knowledge in the first place) and link it to the text. The reading comprehension framework includes general knowledge as one of the complex variables important for successful reading comprehension. The CI model (Kintsch, 1988) also defines these three sources of input from the text, linguistic input, inference made from linguistic input and general background knowledge as essential to forming a coherent message.

Study Evaluation

This study addressed gaps in the literature and built upon previous work by emulating certain aspects (e.g. comparing inferential vs. literal accuracy; inference deficit etc.) and introducing novel domains (i.e. elaborative inference deficit; predictors of inference deficit). In doing so, however, there were necessarily some aspects of previous research that were not modelled. For example, whilst the children with DLD were compared to both TD and LL peers (as children with DLD in school settings are going to be compared with same age peers for academic purposes), there were no language matched controls. As such, we cannot ascertain whether the inferencing skills of the children with DLD were in-line with their language skills. We did consider including a younger, language matched group but then the groups would differ on age and experience. In the current study, the LL group controls for lower language relative to the number of years exposed to academic curriculum better than a language-matched group could. In addition, there is the question regarding which aspects of language should be 'matched'. Language is a multi-faceted construct, and there is no accepted prescription for which aspect of language, or which test(s) of language, is most appropriate for matching groups (Plante, Swisher, Kiernan & Restrepo, 1993). Studies which have compared the inferencing skills of children with DLD relative to younger language ability matched children have differed in terms of the measures used. For example, Bishop and Adams (1992) 'matched' groups on the Test for Reception of Grammar (Bishop, 1983) raw scores, whereas Adams et al. (2009) 'matched' groups based upon the raw scores of Sentence Comprehension based on the ACE 6-11 (Adams, Coke, Crutchley, Hesketh &

Reeves, 2011). It is therefore important for future research to determine whether children with DLD demonstrate inferencing skill in-line with their language proficiency.

Children with low language.

The inclusion of a third LL group facilitated greater insight into where language may be the most important factor and where group membership seemed to predict performance to an extent greater than that of language. While group membership was based on language, the difference in inferencing scores, particularly elaborative deficit scores, seem disproportionately larger than language differences. It is surprising that the DLD were not significantly less successful at lexical inferencing, knowing as we do that children with poorer language skills find learning new vocabulary inferentially more challenging than their TD peers (Cain et al., 2004; Lucas & Norbury, 2017; Nash & Donaldson, 2005), but the percentage of children in the LL group with lexical inferencing deficit was much closer to that of the TD group than the DLD group (please see Table 3 for more details). The greater performance of the LL group relative to the DLD group in response to elaborative inferential questions is not so easily explained. LL was intended to act as a 'midway' group and yet in terms of elaborative inference deficit these children aligned with the TD group, with no significant difference between the two, and was found to be significantly different to the DLD group. When we examined the cohesive inferencing deficit versus the elaborative inferencing deficit of our three groups we found that, like lexical inferencing, and unlike elaborative inferencing, there were no significant differences in cohesive inferencing between the groups. This implies that the DLD group is impaired in the area of elaborative inference beyond their impairment in the area of cohesive and lexical inference. Additionally, they are disproportionately impaired in this field relative to LL and TD peers when compared to literal comprehension and accuracy in response to cohesive and lexical questions.

Impact of background knowledge on inferencing skill.

Beyond language, elaborative inferencing draws upon general world knowledge. Elbro and Buch-Iverson (2013), in an experimental study, piloted a classroom intervention which taught TD children (aged 11-12) how to use background knowledge. They found that only eight 30-minute sessions generated a large training effect on inference skill. A substantial and sustained transfer effect to reading comprehension, not mediated by students' motivation, single word reading, vocabulary or non-verbal IQ was found. By age 10/11, the life experiences of a child with DLD may not be the same as a TD child, or even a LL child. It is known that children with DLD have increased risk of social impairment (Maggio, Grañana, Richaudeau, Torres, Giannotti & Suburo, 2014; Clegg, Hollis, Mawhood & Rutter, 2005; Durkin & Conti-Ramsden, 2007). Conti-Ramsden, Mok, Pickles and Durkin (2013) discuss the difficulties (pragmatic and emotional) that poor communicative skills can create in relating to others, in expressing one's needs or feelings and in understanding messages. Adolescents with a history of DLD have been more likely than their TD peers to report higher levels of peer problems, emotional symptoms, hyperactivity and conduct problems (Conti-Ramsden et al., 2013). It is therefore possible that the group differences in this area may affect life experience to such an extent that it impacts upon the background knowledge a child with DLD will hold, relative to their TD peers. This could explain the comparable inferencing skill of children with DLD and younger children. Zadeh, Im-Bolter and Cohen (2007) posited that children with DLD have an impaired ability to conceptualise the complex and ambiguous worlds of social relationships. Interestingly, this skill may also affect the ability to move beyond the text to one's world knowledge and link this to the text at hand.

Future research aimed at providing a better insight into readers' comprehension monitoring strategies may help to indicate the process by which answers are generated.

Future Research.

The present study has increased the knowledge base regarding the contribution that language and literacy skills make to inferencing competence, but regression models do not yet account for all of the variance. Future research should also therefore explore a greater breadth of variables, such as a measure of life experience and memory (beyond verbal working memory). This may be accomplished using standardised quantitative measures offering an insight into life experience and quality of life (e.g. Kidscreen-27; The Kidscreen Group Europe, 2006), and psychometric measures of memory (e.g. Wide range assessment of memory and learning—second edition; Sheslow & Adams, 2003).

The conclusions about elaborative inferencing study drawn from this study are based on 4 questions in the age-appropriate section of the WIAT-II reading comprehension subtest. In the absence of a standardised measure with more elaborative inferencing questions to substantiate these very interesting findings, one suggestion would be to repeat this study at a different age-group, thus using a different age-normed section and hence different questions. The need for development of a standardised measure specifically targeting elaborative inferencing is discussed further below.

Difficulty with elaborative inferencing could be due to impaired retrieval of appropriate information from text; impaired recollection of background information; impaired integration of new and prior knowledge (Cain et al., 2001). Additionally, a common approach to inferencing categorization could facilitate researchers in this field's ability to build upon prior knowledge and would leave less to interpretation from one researcher to the next. This is not the status quo, as can be seen in Botting and Adams' (2005) and Norbury and Bishop's (2002) use of "bridging inferences" (where new information is related to old, i.e. elaborative inferencing: 'gap-filling') and "logical inferences" (where the relationships between words/referents can be deduced: 'text-connecting'). Furthermore, longitudinal research with greater sample sizes is needed to explore the developmental trajectory of inferencing and to understand how the importance of different predictors may change over time.

As previously discussed, this study found no significant difference between children with LL and TD children in terms of inferencing ability. Therefore, other possible factors influencing reading comprehension in children with LL, such as difficulties with vocabulary acquisition (cf. Cain et al., 2004; Nash & Donaldson, 2005), need to be explored further. These children may not receive the benefit of the full support that a child with a diagnosis of DLD will be entitled to, despite their documented difficulty with reading comprehension (Myers & Botting, 2008). More information is needed to form a standard classroom intervention to prevent an exacerbation of negative outcomes due to LL.

Educational and clinical implications.

As children with DLD find elaborative inferencing disproportionately difficult compared to their TD and LL peers, it is paramount that teaching and learning using this process be rethought when working with children with DLD. Within the collaborative classroom (Hill & Hill, 1990), emphasis is placed upon the learner making their own meaning. Children with DLD may require more guidance during these tasks or require these activities to be more scaffolded. Additionally, in planning assessments of learning, it is important to note that a measure relying on inferencing as a single construct cannot give a true indication of knowledge; children may differ in terms of competency making different types of inferences. On a positive note, the cohesive and lexical inferencing skills of children with DLD were not significantly different to that of their LL or TD peers. These skills may be used to support interventions targeting elaborative inference making skills.

This study also, however, has clinical implications concerning the development of normed assessments of reading comprehension. The WIAT-II purports to measure literal, inferential and lexical knowledge. This is certainly the case, yet these questions are not highlighted as such, hence the need for the researchers in this study to categorise these questions. There is also an unequal number of each type of question present, and only 4 of these measure elaborative inferencing. Given the clear result that children with DLD experience disproportionate difficulty with elaborative inferencing, a measure that mixes elaborative inferential questions with literal questions and other types of inferential questions will only give an overall indication of a child's reading comprehension ability. Measures targeting these domains independently should be established.

Conclusion

To conclude, although children with DLD have poorer literal and inferential reading comprehension than their TD peers, they are likely to find inferential comprehension, especially elaborative inferencing, particularly challenging. It is therefore important that children with DLD are identified and that interventions target those variables found to be predictors of inferencing skill – vocabulary knowledge, single word reading and verbal working memory (cf. Nash & Snowling, 2006). An intervention approach such as that demonstrated by Elbro and Buch-Iverson (2013) could be modified to meet the individual needs of children with DLD. Ideally such support will also be offered to children with poor language skills (but no diagnosis) as they are also at risk for reading comprehension impairments. It is important that the different needs of children with LL are recognised and disparate interventions are developed utilizing the strengths of this group. The effectiveness of such interventions with children with DLD and LL is yet unknown but could improve the outcomes of these children (Snowling & Hulme, 2012).

Acknowledgements

This research was funded by a Vice-Chancellor's Scholarship from XXX. We would like to thank all of the children who took part in this study, as well as their parents and schools; this research would not have been possible without you. Also thank you to XX, who contributed in her role of undergraduate research assistant to data entry and the preparation of this article.

References

Adams, C., Clarke, E., & Haynes, R. (2009). Inference and sentence comprehension in children with specific or pragmatic language impairments. *International Journal of Language & Communication Disorders, 44*(3), 301-318. doi: 10.1080/13682820802051788.

Adams, C., Coke, R., Crutchley, A., Hesketh, A., & Reeves, D. (2001). *Assessment of comprehension and expression 6-11*. NFER-Nelson, London. ISBN 9780708705612

American Psychiatric Association (APA) (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* American Psychiatric Pub.

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*(6), 1173.

Bishop, D. V. M. (1983). *Test for the reception of grammar*. Manchester UK: Manchester University.

Bishop, D. V. M. (1998). Development of the Children’s Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39,* 879–891.

Bishop, D.V.M. (2003a). *Test of reception for grammar - 2 (TROG-2)*. London: Harcourt Assessment.

Bishop, D.V.M. (2003b). *The children's communication checklist – 2 (CCC-2)*. London: The Psychological Corporation.

Bishop, D. V M., & Adams, C. (1992). Comprehension problems in children with Specific Language Impairment: literal and inferential meaning. *Journal of Speech, Language, and Hearing Research, 35*(1), 119-129. doi: 10.1080/13682820601084402.

Bishop, D. V. M., McDonald, D., Bird, S., & Hayiou-Thomas, M. E. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development*, *80*(2), 593-605. doi: 10.1111/j.1467-8624.2009.01281.x.

Bishop, D. V. M., Snowling, M.J., Thompson, P.A., Greenhalgh, T., CATALISE-2 consortium (2016). *CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development. Phase 2. Terminology.* PeerJ Preprints4:e2484v1.

Botting, N., & Adams, C. (2005). Semantic and inferencing abilities in children with communication disorders. *International Journal of Language & Communication Disorders*, *40*(1), 49-66. doi: 10.1080/13682820410001723390.

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. doi: 10.1348/000709904X22674.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., Gong, T., Johns, C. L., & Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of reading. *Reading and Writing*, *29*(3), 435-451. doi: 10.1007/s11145-015-9608-6.

Brooks, G. (2016). *What works for children and young people with literacy difficulties? The effectiveness of intervention schemes*. London, UK: The Dyslexia-SpLD Trust.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*(5-6), 489-503. https://doi.org/10.1023/A:1008084120205.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. https://doi.org/10.3758/BF03196414.

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*(1), 31.

Cartwright, K. B., Marshall, T. R., & Wray, E. (2016). A longitudinal study of the role of reading motivation in primary students' reading comprehension: Implications for a less simple view of reading. *Reading Psychology*, *37*(1), 55-91. doi: 10.1080/02702711.2014.991481.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research*, *45*(6), 1142-1157. doi:10.1044/1092-4388(2002/093).

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders–a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry*, *46*(2), 128-149.

Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of "quick fix" interventions for children with reading disability?. *Scientific Studies of Reading*, *18*(1), 55-73. doi: 10.1080/10888438.2013.836200.

Conti-Ramsden, G., Bishop, D.V.M., Clark, B., Norbury C.F., & Snowling, M.J. (2014). Specific language impairment (SLI): The internet RALLI campaign to raise awareness of SLI. *Psychology of Language and Communication, 18*(2), 143-148. https://doi.org/10.2478/plc-2014-0011.

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. R. (2017). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language and Communication Disorders*.

Conti-Ramsden, G., Mok, P.L.H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioural functioning. *Research in Developmental Disabilities, 34*(11), 4161-4169.

Conti-Ramsden, G., St Clair, M. C., Pickles, A., & Durkin, K. (2012). Developmental trajectories of verbal and nonverbal skills in individuals with a history of specific language impairment: From childhood to adolescence. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1716-1735. doi:10.1044/1092-4388(2012/10-0182)

Crais, E. R., & Chapman, R. S. (1987). Story recall and inferencing skills in language/learning-disabled and nondisabled children. *Journal of Speech and Hearing Disorders*, *52*(1), 50-55. doi:10.1044/jshd.5201.50.

Cromley, J. G. (2009). Reading achievement and science proficiency: International comparisons from the programme on international student assessment. *Reading Psychology*, *30*(2), 89-118. http://dx.doi.org/10.1080/02702710802274903.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the*

International Neuropsychological Society, 15(3), 331-343. https://doi.org/10.1017/S1355617709090481

Dixon P. & Bortolussi, M. (2013). Construction, Integration and Mind Wandering in Reading. *Canadian Journal of Experimental Psychology*. 67(1), 1-10. doi: 10.1037/a0031234

Dockrell, J.E., Lindsay, G.P., Palikara, O., & Cullen, M.A. (2007). *Raising the achievement of children and young people with specific language and communication needs and other special educational needs through secondary school to work and college.* Nottingham: Department for Education and Skills.

Dodwell, K., & Bavin, E. L. (2008). Children with specific language impairment: An investigation of their narratives and memory. *International Journal of Language & Communication Disorders, 43*(2), 201-218. doi: 10.1080/13682820701366147

Durkin, K., & Conti-Ramsden, G. (2007). Language, social behavior, and the quality of friendships in adolescents with and without a history of specific language impairment. *Child Development*, *78*(5), 1441-1457. doi: 10.1111/j.1467-8624.2007.01076.x

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

Ellis Weismer, S. (1985). Constructive comprehension abilities exhibited by language disordered children. *Journal of Speech and Hearing Research*, *28*, 175-184. doi:10.1044/jshr.2802.175

Garnham, A., & Oakhill, J. (1996). The mental models theory of language comprehension. In B.K. Britton & A.C. Graesser (Eds.), *Models of Understanding Text (313-339)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (1-13). NJ: Lawrence Erlbaum Associates.

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

Hernandez, D. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation.

Hill, S., & Hill, T. (1990). *The collaborative classroom: A guide to cooperative learning.* South Yarra, Victoria: Curtain.

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

Karasinski, C., & Weismer, S. E. (2010). Comprehension of inferences in discourse processing by adolescents with and without language impairment. *Journal of Speech, Language, and Hearing Research, 53*(5), 1268-1279. doi: 10.1044/1092-4388(2009/09-0006).

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281-300. doi: 10.1080/10888430802132279.

Kintsch, W. (1988). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychological Review, 95*(2), 163-182.

Kleeck, A. V. (2008). Providing preschool foundations for later reading comprehension: The importance of and ideas for targeting inferencing in storybook-sharing interventions. *Psychology in the Schools, 45*(7), 627-643. doi: 10.1002/pits.20314.

Lucas, R., & Norbury, C. F. (2015). Making inferences from text: it's vocabulary that matters. *Journal of Speech, Language, and Hearing Research*, *58*(4), 1224-1232. doi: 10.1044/2015_JSLHR-L-14-0330.

Lucas, R., & Norbury, C. F. (2017). Can children with Autism Spectrum Disorders learn new vocabulary from linguistic context? *Journal of Autism and Developmental Disorders, 47*(7), 2205-2216. https://doi.org/10.1007/s10803-017-3151-z

Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers’ sentence-and text-level representations. *Journal of Memory and Language*, *36*(1), 129-145. https://doi.org/10.1006/jmla.1996.2485.

Maggio, V., Grañana, N. E., Richaudeau, A., Torres, S., Giannotti, A., & Suburo, A. M. (2014). Behavior problems in children with specific language impairment. *Journal of Child Neurology, 29*(2), 194-202.

McClintock, B., Pesco, D., & Martin-Chang, S. (2014). Thinking aloud: Effects on text comprehension by children with Specific Language Impairment and their peers. *International Journal of Language & Communication Disorders*, *49*(6), 637-648. doi: 10.1111/1460-6984.12081.

Myers, L., & Botting, N. (2008) Literacy in the mainstream inner-city school: Its relationship to spoken language. *Child Language Teaching and Therapy, 24*(1), 95-114.

Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research*, *48*(2), 439-458. doi:10.1044/1092-4388(2005/030).

Nash, H., & Snowling, M. (2006). *Teaching new words to children with poor existing vocabulary knowledge: a controlled evaluation of the definition and context methods. International Journal of Language & Communication Disorders*, *41*(3), 335-354. doi: 10.1080/13682820600602295

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90*, 142–1471. https://doi.org/10.1016/j.jecp.2004.11.003.

Norbury, C. F., & Bishop, D. V. (2002). Inferential processing and story recall in children with communication problems: a comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *International Journal of Language & Communication Disorders*, *37*(3), 227-251. doi: 10.1080/13682820210136269.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *The Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. doi: 10.1111/jcpp.12573.

Norbury, C. F., & Nation, K. (2011). Understanding variability in reading comprehension in adolescents with autism spectrum disorders: Interactions with language status and decoding skill. *Scientific Studies of Reading*, *15*(3), 191-210. http://dx.doi.org/10.1080/10888431003623553.

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.

Palikara, O., Dockrell, J. E., & Lindsay, G. (2011). Patterns of change in the reading decoding and comprehension performance of adolescents with specific language impairment (SLI). *Learning Disabilities: A Contemporary Journal*, *9*(2), 89-105.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading* (pp. 227-247). Oxford: Blackwell Publishing. http://dx.doi.org/10.1002/9780470757642.ch13.

Plante, E., Swisher, L., Kiernan, B., & Restrepo, M. A. (1993). Language matches: Illuminating or confounding?. *Journal of Speech, Language, and Hearing Research*, *36*(4), 772-776. doi:10.1044/jshr.3604.772

Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading comprehension. *Reading and Writing*, *27*(8), 1467-1484. doi: 10.1007/s11145-014-9501-8.

Roth, F. P., Speece, D. L., Cooper, D. H., & Paz, S. D. L. (1996). Unresolved mysteries: How do metalinguistic and narrative skills connect with early reading?. *The Journal of Special Education*, *30*(3), 257-277. doi: 10.1177/002246699603000303.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical evaluation of language fundamentals, fourth edition—Screening test (CELF-4 screening test)*. Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Sheslow, D., & Adams, W. (2003). Wide range assessment of memory and learning—second edition (WRAML2). *Wide range Inc., Delaware*.

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology*, *107*(2), 321. http://dx.doi.org/10.1037/a0037769.

Snowling, M. J., & Hulme, C. (2012). Interventions for children's language and literacy difficulties. *International Journal of Language & Communication Disorders*, *47*(1), 27-34.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E. ,& O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research*, *40*, 1245–1260.

Torgesen J., Wagner R., & Rashotte C. (2011). *Test of word reading efficiency (TOWRE-2).* (2nd ed.). Austin, TX: Pro-Ed.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, *28*(4), 409-426. http://dx.doi.org/10.1080/01443410701708228.

Wechsler, D. (2005). *Wechsler individual achievement test 2nd edition (WIAT II)*. London: The Psychological Corporation.

Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence (WASI-II)*. London: The Psychological Corporation.

Wright, H. H., & Newhoff, M. (2001). Narration abilities of children with language-learning disabilities in response to oral and written stimuli. *American Journal of Speech-Language Pathology*, *10*(3), 308-319. doi:10.1044/1058-0360(2001/027).

Zadeh, Z. Y., Im-Bolter, N., & Cohen, N. J. (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of Abnormal Child Psychology*, *35*(2), 141–152.

Figure Captions

Figure 1. Group differences in correct responses for literal and inferential subtype questions. Error bars represent standard error. The DLD group performed significantly lower on all question types than the TD group (all $p < .020$) but no significant differences were found between the TD or DLD and LL groups (all $p > .439$).

Figure 2. Standardized regression coefficients for the relationship between vocabulary knowledge and inferencing ability as mediated by word reading accuracy. The standardised regression coefficient between vocabulary knowledge and inferencing ability, controlling for word reading, is in parentheses.

Figure 3**.** Percentage of participants in each group with an inferencing deficit relative to the TD mean.

Figure 4. Standardized regression coefficients for the relationship between vocabulary knowledge and elaborative inferencing deficit as mediated by word reading accuracy. The standardised regression coefficient between vocabulary knowledge and elaborative inferencing deficit, controlling for word reading, is in parentheses.