Word Learning from Written Context in Early Childhood

Effects of Book Format, Scaffolding Features, and Child Characteristics

Laura Diprossimo, B.Sc., M.Sc.

This thesis is submitted

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Lancaster University

Department of Psychology

January 2024

Thesis Abstract

The main objective of this thesis is to investigate word learning from written context in early childhood, with a focus on the effects of book format, scaffolding features, and child characteristics. There is consensus that written language provides unique opportunities to foster children's vocabulary development. Yet, early literacy experiences are rapidly changing, with e-books and digital reading environments becoming increasingly common. Prior research investigating the effect of book format on word learning in young children has reported inconsistent findings, as synthesised in a recent meta-analysis. Given the pivotal role of early vocabulary for later reading comprehension, academic achievement, employment, and well-being, it is critical to unpack the effects of book format, scaffolding features, and their interaction with child characteristics in shaping vocabulary acquisition from written context.

Chapter 1 provides an introduction, thesis rationale, and an overview of the literature and theories relevant to the empirical work presented across chapters 2, 3, and 4. Chapter 2 reports a pre-registered study investigating the effects of book format and child characteristics on word learning via naturalistic shared reading. Chapter 3 provides an in-depth evaluation of caregivers' scaffolding during shared reading. The focus of this pre-registered study was to test the generalisability of the fine-tuning hypothesis, that is, whether caregivers adjust their communication to their child's individual lexical knowledge and whether this adjustment supports children's word learning in the context of shared reading. Caregivers' scaffolding behaviour in print- and digital-based shared reading was also compared. Chapter 4 reports a pre-registered study using a big data approach to investigate the associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment. Finally, Chapter 5 discusses the implications of these findings for theory and practice, along with the recommendations and directions for future research.

Table of Contents

List of Tables

Supplementary Tables

List of Figures

Supplementary Figures

Acknowledgements

I complete this thesis full of joy and gratitude. For this, there are many wonderful people to thank. First and foremost, I would like to thank my primary supervisor, Professor Kate Cain. Kate, thank you for being an exceptional mentor and for contributing immeasurably to my growth. I am grateful for your thoughtful guidance and insightful feedback. Thank you for encouraging me to explore research ideas, gain diverse experiences, and prepare for the next steps in my career, all while ensuring that I stayed on track. I would also like to express my gratitude to my secondary supervisor Dr Margriet Groen. I truly enjoyed discussing research ideas and I am grateful for your constructive feedback during the regular appraisals. I would also like to sincerely thank Dr Anastasia Ushakova for being such a generous collaborator and sharing her expertise with all things statistics. Anastasia, every time I approach my analyses, your advice is with me.

A big thank goes to the extended MSCA ITN e-LADDA. Thanks to the PIs who made this possible, particularly, Professor Mila Vulchanova and Dr Valentin Vulchanov, who brought us all together. Mila and Valentin, thank you for hosting me in your department during my secondment at NTNU and for all the fantastic training opportunities during these years. Thank you, Ibrahim El Shemy and Zijian Fan for making me feel at home in Trondheim. I would also like to thank Professor Sue Buckley, for the opportunity to support the organisation of the Online Down Syndrome Research Forum 2021 during my secondment at DSEI. I extend my gratitude to Professors David Saldaña and Isabel Rodriguez Ortiz for hosting me in their department during my secondment at the University of Seville, and for their constructive and insightful feedback at the latest stage of this work. My stay in Seville would not have been the same without the amazing welcome and support of Hülya Aldemir and Ana Lucia Urrea.

Thank you to all the e-LADDA members and particularly to the IP3 team: Martina de Eccher, Paula Janjić, Alireza Kamelabad, Theresa Kalchhauser, Fatih Sivridag, Aisha Futura Tüchler, Ana Lucia Urrea, Karla Zavala Barreda. Working with you has been amazing. Big thanks also go to Luca Raggioli for always having an open ear

every time I wanted to chat about research ideas, career plans, and good food. Thank you, Karla Zavala Barreda, for the many inspiring conversations.

I would also like to thank all the people who made my time at Lancaster University intellectually stimulating and truly enjoyable. Thank you to the amazing members of LULLU, Dr Daphne Barker, Dr Lana Jago, Dr Nicola Currie, Dr Gill Francey, Dr Samantha Russell, Yazhi Lu, and Yawen Ma. Our Lab meetings were a wonderful space to discuss and refine ideas. I must also thank Dr Sayaka Kidby for welcoming me to Lancaster and helping me even before I arrived. I would like to thank Elena Altmann, for the lovely time together, research chats, walks in nature, and continuous support. I must thank Yanxi Lu for many things, but I will just say that Lancaster became a better place after I met you. I would like to thank Drs Francesca Citron and Adina Lewis for bringing Yoga to our department – that made such a positive difference. I am grateful to Dr Elisa Rubegni for the many inspiring conversations and for being such a kind friend. I would like to thank our Lab Manager Dr Katharina Kaduk for making the developmental laboratories the most welcoming and child-friendly places one could imagine. Peter Tovee, Barrie Usherwood, and Stuart Walker, you are the best technical team a researcher could wish for – thank you for your help on many occasions. I would also like to express my gratitude to Dr Jacky Chan for providing incredible support and companionship during our parallel busy testing times. I would like to thank the many children, parents, and teachers, who made this research possible. My gratitude goes to Simya Aravamuthan, Gracey Caller, Arwen Hon, Phoebe Schaw, and Ffion Jones for their fantastic support as student assistants.

Finally, immense credit goes to the wonderful support network that keeps me happy and strong during challenging times, under pressing deadlines, and through transitions between jobs and countries. Thanks to my lifetime friends in Italy who support me from a distance and make me feel at home as soon as we meet, in particular, Anna, Martina, Federica, and Susan. A large credit should go to my family, to whom I would like to dedicate this thesis. Mamma e papà e nonni, siete il mio esempio. Il vostro amore mi dà moltissima forza. Giulia, grazie per sopportare il mio caratterino ed esserci sempre. Víctor, volverte a encontrar ha sido lo más inesperado y lindo de mi vida. Gracias por tu apoyo, tu amor, y tu alegría.

Permission to Submit a Multi-Part Thesis

[Removed as per Lancaster University's guidelines]

Author's Declaration

I declare that this thesis is my own work completed under the primary supervision of Professor Kate Cain. Author contributions are listed at the beginning of each empirical chapter. No portion of the work referred to in this thesis has been submitted elsewhere for another degree or qualification at this or any other institution.

Laura Diprossimo

31.12.2023

Chapter 1: Introduction

Introduction and Thesis Rationale

Word learning rates in early childhood are truly remarkable. It has been estimated that in the first five years of life, children learn on average one new word every one to two waking hours (Tomasello, 2003). This learning rate is even more impressive when we consider its steep growth: children learn on average only one new word per day in the first two years of life, but this learning rate steeply increases towards the end of the preschool period and continues to grow until the school years (Fenson et al., 1994; Tomasello, 2003). For instance, it has been estimated that preschool children know an average of 3,000 words, children in Grade 1 know an average 10,398 words, and children in Grade 3 know an average of 19,412 words (Anglin et al., 1993). Thus, vocabulary development is viewed as an outstanding feat of child development. But this is only one side of the story. Early vocabulary size is also characterised by marked individual differences (Fenson et al., 1994; Frank et al., 2021; Rowe, 2012), and these early differences are stable (Bornstein et al., 2018) and bear a profound and long-term impact on a range of developmental outcomes. For instance, vocabulary at age five is a unique predictor of academic achievement, employment, and well-being (Hoff, 2013; Law et al., 2009; Pace et al., 2019). For these reasons, a central topic in language acquisition is the investigation of the origins and nature of such individual differences.

One important environmental factor associated with differences in early vocabulary is exposure to written language (Mol & Bus, 2011; Sénéchal et al., 2008). Early exposure to written language through shared reading with caregivers has been shown to explain 12% of the variance in pre-schoolers' and kindergarteners' vocabulary (Mol & Bus, 2011). The benefit of exposure to written language can be explained by its statistical properties. The language of children's storybooks contains more unique word types than child-directed language and therefore provides unique opportunities for children's vocabulary growth (Montag et al., 2015; Nation et al., 2022). Written language is an important source for vocabulary development, yet it is increasingly acknowledged that early literacy experiences are rapidly changing. Portable digital media, and consequently digital books, are increasingly accessible to

young children (Hassinger-Das et al., 2020; Ofcom, 2019). These new book presentation formats bring opportunities, but also challenges, for children's word learning (Furenes et al., 2021; Takacs et al., 2015). In this changing landscape, the overarching aim of this thesis is to investigate the factors that influence children's word learning from written context during contemporary reading experiences which involve new media and affordances, while taking into account child characteristics. Below, I review the literature and theory that provide the framework for the three subsequent empirical chapters. In addition to advancing our theoretical understanding of the word learning process in young children, this work is intended to provide critical insights to inform early literacy policies and practice, and maximise learning opportunities for every child in today's digital world.

Literature Review

Word Learning from Written Contexts

Written Contexts. Prior research has estimated that reading a book per day with a child provides exposure to more than a million words by the age of five (J. A. R. Logan et al., 2019). As outlined here, one important motivation to study word learning from written contexts is that written language provides learning opportunities beyond talk (Nation et al., 2022). Several corpus studies have compared the lexical and syntactic properties of written and spoken language (Dawson et al., 2021; Montag, 2019; Montag et al., 2015; Montag & MacDonald, 2015; Nation et al., 2022). When comparing the language of children's books with child-directed language, Montag et al. (2015) found that the former comprised more unique words. Dawson et al. (2021) replicated and extended these findings by showing that the language of children's books is characterised by words that are more abstract, acquired later in development, and more emotionally arousing compared to words in child-directed language. They are also more commonly nouns and adjectives, and they tend to be longer and more morphologically complex. Thus, there is strong evidence that written language substantially extends word learning opportunities for children, as it provides exposure to words that are unlikely to be encountered in spoken language. Written context also provides exposure to rare and complex sentence types, such as passive sentences and sentences containing relative clauses (Montag, 2019; Montag & MacDonald, 2015),

and syntax may, in turn, bootstrap the acquisition of new vocabulary (Rowland, 2013; Syrett et al., 2012; Tomasello, 2003).

Shared Reading. For young children, who have limited language and reading skills, word learning from written contexts happens through shared reading experiences with more knowledgeable adults. During shared reading, adults read the text aloud to the child and scaffold learning through extra-textual talk and non-verbal support – albeit adults shared reading styles can vary (Reese et al., 2003). In the context of shared reading, the following framework has been proposed to examine how young children process book information to extract word meanings (Sénéchal et al., 1995; Sénéchal & Cornell, 1993). First, children need to encode and maintain in memory a phonological representation of the novel word; second, they need to extract clues from the semantic, syntactic, and pictorial contexts to constrain memory search for potential meanings in the case of learning synonyms for known referents, and to facilitate the inferential process in the case of novel referents; third, children need to select and construct an appropriate meaning; then, they need to associate the inferred meaning with the phonological representation of the novel word; and finally children need to integrate and store the new knowledge with their existing knowledge base. This framework is particularly useful in making explicit the core processes involved in word learning from shared reading. It immediately becomes apparent that cognitive skills, including memory, attention, and prior knowledge, are crucial for children to learn new words from shared reading. I will return to this point when discussing the role of child characteristics in more detail in subsequent sections. Another theoretical framework that can usefully inform the study and interpretation of word learning from shared book reading is the dual-coding theory (Paivio & Csapo, 1973). According to the dual coding theory, image and verbal codes are independent and have an additive effect on recall and learning. In the context of shared reading, where children are read to while they are shown corresponding pictures through the storybook, the combination of image and verbal codes may be particularly helpful to support their learning of word forms and meanings.

Independent Reading. Word learning from written context in early childhood goes beyond shared reading. As schooling begins and children develop their reading skills, independent reading becomes an important source of word learning (Cain, 2007;

Cain, Lemmon, et al., 2004; Tomasello, 2003). Word learning from independent reading differs substantially from shared reading. One notable difference is the reduction of multimodal supports in books designed for independent readers, where pictorial information is reduced or absent. Similarly, the multimodal scaffolding provided by caregivers is typically not available during independent reading. Therefore, the inferential load on the child becomes greater. Proficiency in reading skills, particularly reading comprehension, becomes critical to efficiently derive novel word meanings from the surrounding linguistic context (Cain et al., 2003; Cain, Lemmon, et al., 2004). Thus, well-designed digital media offer the promise to provide timely multimodal scaffolds and support word learning during independent reading for beginner readers.

Book Format

Contemporary reading experiences involve new forms of reading on media such as iPads and tablets (Kucirkova, 2019). This has opened the debate on the effectiveness of these new book formats, relative to traditional print books. This debate can be situated within the broader context of the *children and screens* discussion (M. L. Allen et al., 2016; Bus et al., 2020; Hassinger-Das et al., 2020; Vulchanova et al., 2017). Children are seen as part of a natural experiment, surrounded by portable and easy-to-access digital media. Concerns about the effect of screens on child development have been raised since the introduction and popularity of television and computers, but the advent of iPads in 2010 has revived the discussion, since portable media can be used by children everywhere and anytime with much more flexibility (Hassinger-Das et al., 2020).

There is increasing consensus that screen time is a poorly specified concept and that digital content and the context of (co-)use are important factors to consider when studying the effect of digital media on learning and development (Hassinger-Das et al., 2020; Madigan et al., 2020). Accordingly, the *displacement hypothesis*, referring to the idea that interactions with digital media would interfere with other important activities for child development, including shared reading, is being reconsidered. Well-designed digital media are increasingly seen as extending educational opportunities and having the potential to enrich children's learning environments (Bus et al., 2020; Kucirkova, 2019). Yet more investigation, as well as partnerships between developmental

researchers and software developers, are needed to make this happen (Diprossimo et al., 2023; Hassinger-Das et al., 2020).

In this context, one of the research foci of this thesis was to compare word learning from print and digital books during shared reading. A relevant meta-analysis was published by another research group as this doctoral research was developed: this meta-analysis examined the effect of book format on word learning in children up to 8 years of age and found no clear difference between print and digital format (Furenes et al., 2021). Such null findings do not have a straightforward interpretation, as is always the case in the framework of null hypothesis significance testing, and even more so when we consider that many studies are difficult to interpret as they lack the power to detect any meaningful effect due to small sample sizes. The inconsistent findings amongst studies invite further replications with sufficiently powered samples. This issue is addressed in the empirical study reported in Chapter 2.

Scaffolding Features

Caregiver Scaffolding. The multimodal nature of shared reading experiences is substantially extended by caregivers' mediation. Caregivers provide verbal and non-verbal support during shared reading, yet the reading style differs substantially between caregivers (Reese et al., 2003). In line with the notion of *scaffolding* (Wood et al., 1976) and *zone of proximal development* (Vygotsky, 1978), adults' support facilitates children's word learning from shared reading (for a meta-analysis see Flack et al., 2018). Prior research suggests that verbal scaffolds, including repetitions, definitions, and questions in extra-textual talk support children's word learning from shared book reading (Blewitt et al., 2009; Flack et al., 2018; Lenhart et al., 2019). Gestures, including pointing and iconic gestures, have also been shown to support word learning from shared reading (Barnes et al., 2023; Flack et al., 2018; Flack & Horst, 2018),

When considering the framework proposed by Sénéchal et al. (1995), it becomes apparent that pointing reduces the inferential load associated with identifying a new referent in the pictorial context by directing children's attention; and repetitions and questions assist children in associating the new referent with its label. In line with the dual-coding theory (Paivio & Csapo, 1973), pointing may boost word learning by focusing attention on relevant pictorial information. Representational and

iconic gestures can support word learning by providing multimodal cues for word meanings. This is consistent with theories of embodied cognition (e.g., Sadoski & Lawrence, 2023; Smith & Gasser, 2005), which posit that mental representations, including word meanings, are grounded in sensory and sensorimotor experience.

Research on child-directed language in contexts different from shared reading can provide insights and help generate new hypotheses about the features of caregivers' communication that support word learning in the context of shared reading. Here, of particular interest is the so-called fine-tuning hypothesis, according to which, caregivers adapt their communication to their children's individual lexical knowledge in a way that supports their word learning (Leung et al., 2021). This hypothesis reflects an increasing interest in understanding the dynamic and interactive mechanisms that support children's word learning and language development more broadly (Donnellan et al., 2023; Fusaroli et al., 2023; Leung et al., 2021; Shi et al., 2022). This recent work can be situated within a relatively long tradition of research documenting that caregivers adapt their language when talking to children compared to adults (Fernald & Mazzie, 1991; Hills, 2013; Newport & Gleitman, 1984; Onnis et al., 2008), and that the features of child-directed language change as children age (Huttenlocher et al., 2010). The important extension proposed by the fine-tuning hypothesis is that caregivers not only adapt their communication to the general linguistic level of their child, but also modulate it at a more fine-grained level as a function of their child's knowledge of specific words. This modulation is expected to be particularly beneficial for children as it would provide just the right level of support to learn unknown words, in line with the scaffolding principle (Wood et al., 1976) and the notion of the zone of proximal development (Vygotsky, 1978). The generalisability of the finetuning hypothesis to the context of shared reading is tested in the study reported in Chapter 3.

It is well established that word learning from shared reading is supported by adult mediation through verbal and non-verbal scaffolds. How book presentation format influences caregivers' behaviours remains an open question. Prior research suggests that the book presentation format influences the quality of interactions during shared reading (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020). For instance, digital books are associated with less communicative initiations, responses

and expanding talk by mothers (Korat & Or, 2010), and with lower social reciprocity between the parent-toddler dyads (Munzer et al., 2019). Chapter 3 reports a study exploring if, and how, book presentation format influences caregivers' scaffolding.

Digital Scaffolding. In addition to comparing the effects of print and digital books that only differ by digitation, the interactive affordances of digital media were also considered, as recommended by the third space theory (Kucirkova, 2017). According to this theoretical framework, the difference between media needs to be embraced rather than neglected (Kucirkova, 2017). It is important to consider that digital media and, consequently, digital texts afford novel opportunities to scaffold children's understanding and learning (Dalton et al., 2011; Gonzalez, 2014; Proctor et al., 2007; Zou et al., 2021). For instance, multimodal support for the pronunciation and meaning of words can be provided through built-in scaffolds in digital texts (Dalton et al., 2011; Furenes et al., 2021; Gonzalez, 2014; Proctor et al., 2007; Yun, 2011). Despite the promise of these features, the evidence of their efficacy is mixed (Abraham, 2008; Furenes et al., 2021). This points to the importance of carefully investigating the factors that mediate their effectiveness. The cognitive theory of multimedia learning (Mayer, 2003) predicts that learning is optimised by well-designed materials consisting of words and pictures compared to words alone, and provides guiding principles for the design of multimedia material. Those principles include the *multimedia effect*, according to which learning is deeper with words and pictures than from words alone; the *coherence effect*, according to which learning is enhanced when extraneous material is excluded; and the *spatial contiguity effect*, which refers to the learning advantage derived from presenting printed words near rather than far from the corresponding images. Chapter 4 reports on a study investigating the uptake and effectiveness of multimodal scaffolding features to shed further clarity on the conditions under which built-in multimodal scaffolds are most effective.

Child Characteristics

Child characteristics are viewed as an important factor to consider when seeking to understand individual variation in language learning (Rowland, 2013). A careful consideration of learners' characteristics is also critical when studying word learning from different media. Both the capacity model (Fisch, 2000) and the differential susceptibility to media effects model (Valkenburg & Peter, 2013)

underscore the importance of considering learners' characteristics. The capacity model was developed to explain the mechanisms underlying children's comprehension of educational content on television. It was informed by empirical evidence showing that the comprehension of television content is predicted by the amount of invested mental effort (Salomon, 1984). The capacity model highlights the critical role of children's prior knowledge and working memory capacity in predicting comprehension and learning. The differential susceptibility to media effects model builds on prior microlevel media-effects theories, devoting particular attention to the individual media user. This model also emphasises the role of individual characteristics by assuming that media effects are *conditional* on certain user characteristics that act as moderators. Against this theoretical background, a common thread between the studies reported in the present thesis is the focus on child characteristics.

There is evidence to suggest that prior vocabulary knowledge predicts children's word learning from storybooks (e.g., Sénéchal et al., 1995). One possible explanation for the role of prior vocabulary knowledge is that the existing knowledge base and category knowledge facilitate the acquisition of new words (Borovsky et al., 2016). In the context of shared reading, greater vocabulary knowledge also facilitates the understanding of the storyline, freeing up cognitive resources to encode new words as they appear.

Visual attention has also been linked to word learning from shared book reading (O'Toole & Kannass, 2018). Visual attention has long been used as a proxy for engagement, therefore its effect on learning could be explained by enhanced engagement with the learning materials. Working memory, inhibitory control, and executive functions have been shown to predict children's word learning in experimental settings (Baddeley et al., 1998; Gathercole et al., 1997, 1999; Kapa & Erikson, 2020), as well as in the shared reading context (Hadley et al., 2021). The role of working memory, particularly the phonological loop, is central to learning the phonological form of new words (Baddeley et al., 1998), as it provides a mechanism for the temporary storage of new phonological forms whilst more stable long-term representations are being formed. Inhibitory control is associated with word learning as it enables learners to suppress irrelevant targets and focus on relevant ones (Kapa & Erikson, 2020). Age and gender may also play a role in word learning from shared book

reading (Bergman Deitcher et al., 2019; Reich et al., 2019). Importantly, prior research underscores the importance of considering a range of child characteristics that may interact with the book format or scaffolding features to influence understanding and learning (Eng et al., 2019; Salmerón et al., 2021; Takacs et al., 2015).

Scaffolding features embedded in digital media can be made available to learners on an as-needed basis. In this scenario, the uptake of scaffolds relies on learners' metacognitive awareness. Such metacognitive ability is likely to vary as a function of child characteristics, including ability level and age (Baker & Cerro, 2000; Kirby & Moore, 1987), gender (Sadeghi & Khezrlou, 2012; Tseng et al., 2006; Wu, 2014), and degree of bilingualism (Abu Rabia, 2019). Child characteristics may also modulate the effectiveness of scaffolds (Yun, 2011). For instance, scaffolds may be more effective for beginner compared to advanced learners.

Child characteristics of age, gender, and verbal/reading ability were considered in each of the experimental chapters and, additionally, working memory, visual attention, and executive functions were considered in Chapter 2.

Interim Conclusions

This literature review highlights the important role of written language for children's vocabulary development. Thus, understanding the factors that influence word learning from written context is essential, particularly, in the current digital ecology, where new reading media are increasingly common. The effects of book formats, scaffolding features, and their interaction with child characteristics are still poorly understood. This thesis aims to fill these gaps in our knowledge and provide further clarity on the complex interplay between child characteristics and new reading media and their affordances.

Thesis Objectives

Given the background discussed above, there were three aims of this thesis, each a focus of the three experimental chapters that follow. The first aim was to examine the effect of book presentation format (print vs. digital) on 4- to 5-year-old children's word learning in the context of shared reading, addressing the contradictory findings of prior research (Chapter 2). The second aim was to examine caregivers' scaffolding during shared reading with print and digital books and test the generalisability of the fine-tuning hypothesis, that is, the extent to which caregivers

tailor input to their child's individual lexical knowledge to support their word learning (Chapter 3). The third and last goal was to explore the association between the use of vocabulary scaffolding features embedded in a digital reading environment and its relation to 5- to 8-year-olds' reading comprehension (Chapter 4).

Methodological Considerations

From a methodological standpoint, the studies reported below are characterised by high levels of ecological validity (Kihlstrom, 2021). Generalising these findings to real-life settings outside the laboratory is critical given their implications for early literacy policy and practice. To achieve this goal, different methodologies were employed. First, a semi-naturalistic paradigm where the caregiver, not the researcher, read the story to the child was developed and implemented in Chapters 2 and 3. By carefully controlling the reading material, book format and setting, a trade-off between experimental control and ecological validity was achieved. A novel big-data approach was implemented in Chapter 4 and extended the ecological validity of this work further. A big dataset of log files resulting from children's interaction with a digital reading supplement in the classroom was analysed. This enabled the study of both the uptake and effectiveness of vocabulary scaffolds in real-life settings. Another important methodological choice that needs to be highlighted here is the inclusion and consideration, where possible, of different word types or word learning measures. I will return to this point in the General Discussion (Chapter 5).

Chapter 2: Word Learning via Naturalistic Shared Reading: Effects of Book Format (Print vs Digital) and Child Characteristics

Linking Statement

In this pre-registered study (OSF Project https://osf.io/6uem9/; Registration https://doi.org/10.17605/OSF.IO/ANCSX), the effect of book presentation format (print vs. digital) on 4- to 5-year-old children's word learning in the context of shared reading was examined. This empirical chapter is foundational as it addresses the contradictory findings of prior research synthesised in recent meta-analyses (Furenes et al., 2021). To do so, a within-subjects design was used to minimise variation between participants, and a large sample of dyads (N = 100) was recruited to increase our statistical power relative to previous research. Further, having the same reader narrating the text across book formats enabled to rule out potential confounds identified in previous research. Finally, a semi-naturalistic paradigm was developed to support the generalisation of findings outside the laboratory.

Author Contribution Statement

In the chapter entitled “Word Learning via Naturalistic Shared Reading: Effects of Book Format (Print vs Digital) and Child Characteristics” the authors agree to the following contributions:

Laura Diprossimo, 80% (Conceptualization, Formal analysis, Data curation, Writing – original draft, Visualization, Project administration)

Kate Cain, 20% (Conceptualization, Writing – original draft, Supervision, Funding acquisition)

Abstract

Shared book reading provides unique opportunities to foster children's vocabulary development. Yet the factors contributing to word learning via naturalistic shared reading are still poorly understood, particularly, in relation to the effects of new media of story presentation. The present study examined the influence of book format (print vs digital) on young children's word learning via naturalistic shared reading, with a focus on the role of individual differences in prior vocabulary knowledge and executive function. British English-speaking caregivers and their 4- to 5-year-olds took part in this pre-registered study (N_{dyads} = 100; child's M_{age} = 57.7 months; 57% girls). In a cross-sectional, within-subjects design, dyads read one of two custom storybooks in print, and the other one in digital format, with order of presentation counterbalanced across participants. Word learning was assessed with phonological recall, semantic recall, and semantic recognition tasks. There was no evidence of a main effect of book format across word learning measures. However, for semantic recognition, there was a significant cross-over interaction between book format and executive function: digital format was beneficial for children with higher executive function, and detrimental for children with lower executive function. Greater prior vocabulary knowledge led to better word learning, and boys were significantly more accurate than girls, across word learning tasks. Further, executive function significantly predicted semantic recall. Overall, these findings underscore the importance of considering individual differences and multiple outcomes when studying learning from different media.

Keywords: word learning, shared book reading, digital media, vocabulary, executive function, individual differences

Public Significance Statement

This study compared 4- to 5-year-olds' word learning from print and digital books during shared reading with a caregiver. Cognitive skills that support attention and memory were more important for learning new words from digital than print books. Prior vocabulary knowledge was equally important for word learning from both media. These results demonstrate the importance of considering child characteristics to understand what influences learning from different media.

Introduction

Developing breadth and depth of vocabulary in early childhood is vital for later reading comprehension, academic achievement, employment, and well-being (Hoff, 2013; Law et al., 2009; Pace et al., 2019). Individual differences in early vocabulary are substantial (Rowe et al., 2012), demonstrating the importance of understanding which factors influence word learning. An important context for early word learning is shared book reading (Dowdall et al., 2020; Elley, 1989; Farrant & Zubrick, 2012). It has been estimated that a book per day provides exposure to approximately 1.4 million words by the age of five (J. A. R. Logan et al., 2019). Moreover, it has been documented that the language of storybooks is lexically and syntactically richer compared to spoken language (Montag, 2019; Montag et al., 2015; Montag & MacDonald, 2015). Therefore, exposure to storybooks, not only increases the quantity of exposure to language, but also the quality. This is particularly relevant since longitudinal work has indicated that input quality, such as the use of diverse and sophisticated vocabulary and decontextualized language with toddlers and pre-schoolers, explains unique variance in their vocabulary one year later (Rowe, 2012).

Shared book reading episodes very often go beyond the verbatim reading of the text to the child. When reading a story to a child, caregivers can engage in a conversation, pose questions, relate the story content to the child's experience, provide repetitions, point to target objects in the storybook, and/or recast the child's production. This set of tutorial behaviours has been referred to as *dialogic reading* and has been shown to support word learning (Flack et al., 2018; Price & Kalil, 2019; Whitehurst et al., 1988). For instance, asking the child questions of increasing complexity about target words has been shown to facilitate the acquisition of elaborated word meaning (Blewitt et al., 2009). As such, the social interaction unfolding during shared book-reading episodes may be an important driver of word learning in naturalistic settings.

Whilst much evidence underscores the critical role of shared book reading for early vocabulary development, a meta-analysis quantifying the effect of shared reading interventions on children's language skills suggests limited effectiveness (Noble et al., 2019). As a result, recent work has pointed to the importance of understanding which factors contribute to word learning from shared reading to

inform and optimize the design of future interventions (Hadley et al., 2021). We review the literature on the effect of book format and child characteristics on word learning from shared reading, which are key factors considered in the present study.

Effects of Book Format

The emergence of new media for story presentation, such as iPads and tablets, has raised questions as to whether these new book formats are as effective as traditional print books in supporting vocabulary acquisition. Prior research has reported inconsistent findings with respect to the presence and direction of the effect. A meta-analysis examining the effect of book format on story comprehension and word learning in children up to 8 years of age, found lower story comprehension for screen compared to print (i.e., screen inferiority effect), which is in line with meta-analytic evidence on adult readers (Delgado et al., 2018; Salmerón et al., 2023). The results for word learning were inconclusive (Furenes et al., 2021). Specifically, the digital format had a significant and positive effect on word learning but, when the authors imputed studies with small sample sizes due to asymmetry around the point estimate, the effect size reduced and was no longer significantly different from zero. This suggests no clear difference between print and digital format. Such null findings do not have a straightforward interpretation in the context of null hypothesis significance testing because the absence of evidence cannot be taken as evidence of absence.

The findings of many studies are difficult to interpret, because they may lack the power to detect any meaningful effect; small sample sizes and between-subject designs are quite common across studies looking at the effects of book format. Another emergent issue affecting the interpretation of some previous studies arises from confounds in the experimental design; for example, in some studies, the reader of the print book (parent) differs from that of the digital book (voiceover). In such scenarios, it is not possible to tease apart any effect of the medium from that of the reader. Further, the majority of work comparing print and digital format to date has focused on a narrow sample of the world's population, with most studies being conducted in the US, Canada, Israel, and the Netherlands (Furenes et al., 2021). Given that experiences with digital media are likely to differ across cultures, it is important to replicate findings in different cultural contexts.

Effects of Child Characteristics

Child characteristics have not always been taken into account in prior research, but they play an important role in word learning and, more broadly, in learning from different media (e.g., Fisher, 2000; Valkenburg & Peter, 2013). For instance, prior vocabulary knowledge significantly predicts young children's word learning from storybooks (Sénéchal et al., 1995) and visual attention has also been linked to word learning from shared book reading (O'Toole & Kannass, 2018). Working memory, inhibitory control, and executive functions predict young children's word learning in experimental tasks (Baddeley et al., 1998; Gathercole et al., 1997, 1999; Kapa & Erikson, 2020). There has been less work exploring their role in word learning from shared reading (Hadley et al., 2021), but a consideration of these factors is critical to provide an accurate estimate of any unique or mediated effect of book format. How such child characteristics potentially interact with book presentation format to influence word learning was a focus of this study.

The evidence on the interplay between book format and child characteristics to date is scant and inconclusive. There may be benefits from a digital learning environment: Multimodal scaffolding features in digital reading are more effective for children with, or at risk for, low literacy skills (Diprossimo et al., 2023; Takacs et al., 2015) or less developed attention regulation (Eng et al., 2019). However, poorer reading comprehension is found for reading on screen compared to print and reported to be more pronounced in children with lower reading skills (Salmerón et al., 2021). Thus, to understand any effect of book presentation format on word learning from shared reading, it is essential to consider a range of child characteristics that may interact with the format and influence learning. Age and gender may also play a role in word learning from shared book reading (Bergman Deitcher et al., 2019; Reich et al., 2019) so were controlled in our analyses to provide unbiased estimates of any unique or mediated effect of book format.

Methodological Considerations

Much previous research has typically considered a single dimension of word learning by assessing semantic recognition with a classic four alternative forced-choice task. Yet, lexical representations of expressive vocabulary can include phonological, semantic, and morphological information (Hadley et al., 2016; Hadley & Dickinson,

2020). An innovative feature of our study is the inclusion of different word learning measures that tap into both phonological and semantic representations. Particularly, definition or semantic recall tasks are used to assess the acquisition of word meaning along a continuum, thus providing a more fine-grain measure of word learning. Another important aspect to consider when studying the effect of book format on word learning is the trade-off between ecological validity and experimental control. Previous research has looked either at the effect of book format on the quality of interactions during shared reading in naturalistic settings (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020) or focused on word learning outcomes in more controlled settings, such as when an experimenter reads the story to the child (O'Toole & Kannass, 2018). There is a well-documented role of adult mediation on children's word learning from storybooks (Blewitt et al., 2009; Blewitt & Langan, 2016), and also the critical need to generalise these results to the real world to inform recommendations and guidelines for users of research (i.e., stakeholders such as parents and early years practitioners). For those reasons, we developed a paradigm where we carefully controlled what caregivers and children read and the format in which they read, but not *how* they read, thereby increasing the likelihood that our findings will generalise to real-life situations.

The Present Study

The present study was designed to address the critical gaps in our knowledge about the effect of book format on young children's word learning from shared reading through the following research questions:

1. Does book format (print vs. digital) influence young children's world learning from naturalistic shared reading?
2. Does book format interact with child characteristics (i.e., prior vocabulary knowledge and executive function) to influence word learning from naturalist shared reading?

Previous research has reported contradictory findings (Furenes et al., 2021). If the screen inferiority effect (Delgado et al., 2018; Furenes et al., 2021) extends to word learning measures, there should be detrimental performance with the digital shared reading format. Previous research indicates that the effects of book format are conditional on a range of child characteristics (Diprossimo et al., 2023; Eng et al., 2019;

Salmerón et al., 2021; Takacs et al., 2015). We explored this possibility in our analyses. One mechanism through which book format may influence word learning is by differently taxing children's cognitive resources. Handling and operating a digital device, albeit with adult support, and the potential novelty of the device itself, may distract children from the content of the story including the new vocabulary it introduces. This may have a detrimental effect on word learning, particularly in children with less developed executive function skills.

Methods

Participants

British English-speaking caregiver-child dyads took part in this study. The sample size of N_{dyads} = 100 was decided a priori informed by previous high-powered studies (Reich et al., 2019) and the use of a within-subject design to further increase the power to detect a statistically significant and meaningful effect. Children were typically developing, as reported by their caregivers, and aged between 48 to 71 months (M_{age} = 57.68; SD = 7.24; 57 girls). Participating caregivers were predominantly highly educated individuals, with 75% achieving an undergraduate degree or higher. Caregivers were aged 29 to 47 years (M_{age} = 37.76; SD = 3.93; 83 self-reported as women; 6 as men; 11 did not state). Eight additional participants were tested but excluded due to missing data on key variables (i.e., memory and word learning measures), following the protocol outlined in our pre-registration. According to caregivers' reports, most of the children in our sample engaged in print-based shared reading activities daily (86%). A large proportion of children never engaged in digital-based shared reading (68%). A detailed breakdown of children's experiences with print and digital books, alone and in the shared reading context, and additional socio-demographics are reported in Table S1 in the Supplementary Materials.

Participants were recruited via the university database, social media advertisements, and flyers distributed in public book libraries in a middle-sized town in the North West of England. This was complemented by a snowball strategy, where participating caregivers were asked to share the study flyer their own social networks. Written informed consent was obtained from caregivers prior to data collection. Children received a book and caregivers received a travel reimbursement for their participation. This research has received ethical approval from the Ethics Committee of

the Faculty of Science and Technology, Lancaster University (reference number: FST-2023-0791-SA-1).

Measures

Vocabulary

Children completed a standardised vocabulary assessment: the Word Classes subtest of the Clinical Evaluation of Language Fundamentals – Preschool-2 (CELF-P2; Wiig et al., 2004). Children are shown three to four pictures (e.g., apple, shoe, bread), while the experimenter names each picture aloud. Children are instructed to identify the two words that go together best (e.g., apple and bread) and explain how they go together (e.g., both are food). Children get 1 point for each correct response, summed to provide a receptive, expressive, and total vocabulary score. This assessment was selected to go beyond single recognition to capture the depth of prior vocabulary knowledge.

Attention, Memory, and Executive Function

Children completed the visual attention subtests of the Developmental Neuropsychological Assessments (NEPSY; Korkman et al., 1998). These subtests are designed to assess the speed and accuracy with which a child can focus selectively on and maintain attention to visual targets (e.g., animals, faces). Children are asked to mark with a crayon, as quickly and as accurately as possible, visual targets across two arrays containing both targets and distractors. An accuracy score is calculated by subtracting the number of non-target marked (commission errors) from the target marked correctly. The total score is determined using the accuracy score in conjunction with time taken to complete the task (speed). Different materials are used for different age groups, with difficulty increasing with age. For this reason, scaled (not raw) scores are reported and used in the analyses.

Two researcher-developed measures of verbal working memory were administered based on the format of the measures used by Gathercole and colleagues (1992). In the forward digit span task, children were asked to repeat unique strings of numbers exactly in the same order as they heard them, starting from one digit. There were three trials at each level of difficulty, and testing ceased when the child got two out of three trials wrong. In the backwards digit span, children were instructed to repeat unique strings of numbers backwards. The same stopping rule was applied.

There were 2 practice trials to model behaviour before the test trials. Children got 1 point for correctly repeated trials. The digit span has been used successfully in prior research with 4- to 5-year-olds and has appropriate test-retest reliability (r = .77, df = 63, p < .001) (Gathercole et al., 1997).

Caregivers completed the Behaviour Rating Inventory of Executive Function, Preschool Version (BRIEF-P; Gioia et al., 2003), which provides a measure of executive functions (inhibition, working memory, shifting, planning/organizing, and emotional control), in ecologically valid settings.

Storybook materials

Custom storybooks were developed to ensure that the storyline would be unfamiliar to all participants and that all target words would be of comparable salience in the storyline. This enabled us to carefully control the size of the book across formats (single page size: 126 x 113 mm; open book / iPad screen size: 126 x 226 mm). Two storybooks with embedded low-frequency words were created in print and digital format so that each book served as the digital condition for half of the participating dyads, and the print condition for the other half. These are available on OSF (https://osf.io/6uem9/?view_only=7bb996db9f7c4d54a7b3bf85723a0b16) under Creative Commons Attribution 4.0 International. Our plots featured a canonical Western structure of exposition, conflict, and resolution.

In experimental settings, word learning is often measured using pseudo-words paired with novel objects. However, this approach did not represent a viable option because, in our study, the caregivers read aloud the book to the children; caregivers would struggle to read and explain pseudowords without significant prior training, limiting the ecological validity of our task. To measure word learning we, therefore, identified real words that were unlikely to be known by children in our age range. Selection involve a range of criteria: their frequency in the SUBLEX corpus of children's tv programs (van Heuven et al., 2014); their age of acquisition (Kuperman et al., 2012); and by examination of the high-difficulty items in standardised vocabulary assessments, such as the British Picture Vocabulary Scale (BPVS; Dunn & Dunn, 2009) (see Lenhart et al., 2020 and Sarı et al., 2019 for a similar approach). The following animal names: *myna, okapi, sloth,* and tools*: clamp, valve, and chisel* were selected. We also included one word in each category more likely to be known by children (e.g.,

toucan, screw) to support motivation and engagement with the storyline. We embedded three target words in each story. Each word was repeated three times across two successive pages (informed by Flack et al., 2018), and accompanied by an adjective describing a visible property on one occasion in the visual and scripted storyline. This was to promote the encoding of semantic features and reflect the rich semantic context in which words typically appear in storybooks. Psycholinguistic properties of target words and accompanying adjectives are reported in Table 1.

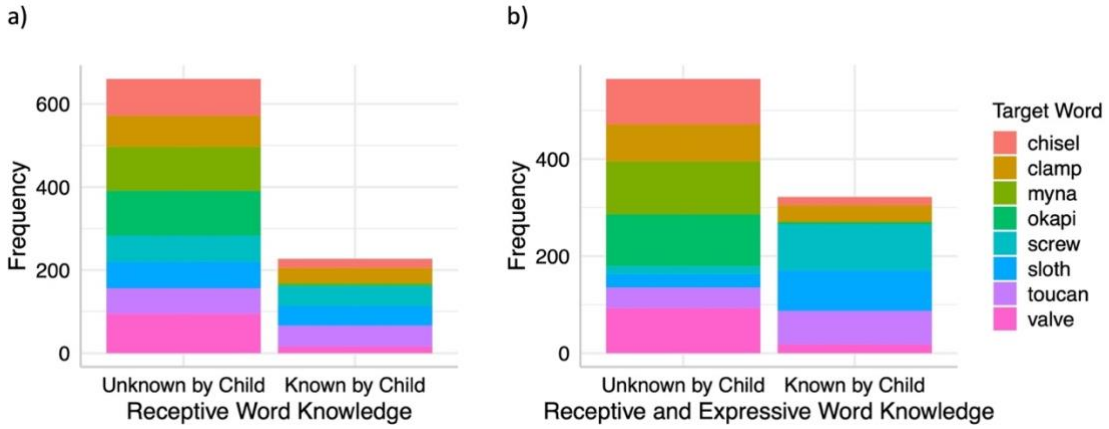
Table 1 *Frequency and Age of Acquisition of Target Words and Accompanying Adjectives*

	Story A			Story B	
	Frequency	AoA		Frequency	AoA
Target words					
okapi	2.23	11.22	clamp	2.71	10.89
sloth	3.38	8.37	valve	3.31	10.78
myna	2.23	9.5	chisel	2.93	10.53
toucan	3.71	8.69	screw	3.82	6.65
Adjectives					
striped	3.08	4.65	wooden	4.34	5.89
furry	4.62	5.72	shiny	4.32	5.05
dark	5.28	3.74	sharp	4.82	6.11
colourful	4.82	4.89	pointy	4.63	7.39

Note. Frequency of SUBLEX. Values lower than 3.5 are considered infrequent in the corpus. AoA = Age of Acquisition.

The selection of target words was further validated via a checklist designed after The MacArthur Communicative Development Inventories (CDI; Fenson et al., 1993). Before the shared reading activities, caregivers completed the vocabulary checklist as a proxy for their child's receptive and expressive knowledge of each target word (see Shi et al., 2022 for similar approach). Caregivers' responses are depicted in Figure 1 and show that the selected target words were largely unknown by children in our age range.

Figure 1 Caregivers' Ratings of their Child Target Word Knowledge



Note. Caregivers' ratings of their child a) receptive word knowledge and b) receptive and expressive word knowledge by target word, colour coded.

Word learning assessments

Phonological and semantic aspects of word learning were assessed after the shared reading episodes. A five-minute delay was introduced to assess retention (in line with Hartley et al., 2020). Phonological recall was measured with a picture naming task (Adlof et al., 2021; Blewitt et al., 2009; Hargrave & Sénéchal, 2000). This word-learning measure taps into children's ability to recall the phonological form of a target word upon presentation of the corresponding visual depiction. As such, this task probes children's learning of the association between the phonological form of a word and its meaning. Pictures of target words were shown either on screen or on paper of identical size (126 x 226 mm), congruent with the book format (print or digital) in which the words were introduced during the shared reading. The researcher then said, "Tell me what this is called". Similar to Blewitt et al., (2009) the pictures were different from those in the books, this was done to assess children's generalisation to different pictures, and thus learning of word meanings beyond the association between a phonological form and a specific picture. Familiar words (dog, cat; pencil, spoon) were interspersed among the targets to maximise the opportunity for children to experience success during testing. Each correct response for a target word was assigned 1 point.

To assess semantic recall, children were asked to describe what they knew about a word. This word-learning measure taps into children's deeper representation of word meaning, and learning the association between the phonological form of a word and its meaning. This task also requires the metalinguistic ability to reflect about one's language knowledge. We adopted the child-friendly procedure used in previous research (Blewitt & Langan, 2016). Children were introduced to a stuffed animal named ''Toby." and told that ''Toby does not know very many words.". For each target word, children were asked ''Can you tell Toby what a [target] is?" After their initial responses, follow-up prompts were given such as ''What else can you tell Toby about a [target]?" and continued until children could add no more information. Children were familiarised with the task via practice trials with familiar words (e.g., dog, cat, spoon, pencil). For each word, children receive 1 point for each unit of relevant information provided (e.g., object function, physical properties). Inter-rater reliability was computed via an intra-class correlation (ICC) analysis with absolute agreement (ICC = .95) (representing excellent agreement, Cicchetti, 1994).

To assess semantic recognition, children were asked to identify referents of the target words in a multiple-choice task (see Blewitt et al. 2009 and O'Toole & Kannass (2018 for a similar approach). This word-learning measure taps into children's learning of the association between the phonological form of a word and its meaning. This task is less demanding compared to those presented above as it does not require a verbal response, but simpler recognition via pointing. Children were asked to ''Find the [target, e.g., clamp]" on a page of four pictures. Each page depicted the target item and three distractors: an item from the same superordinate category as the target, an item from the story, and an item from the same superordinate category as another story word. It is worth noting that children can apply different strategies to identify the referent, including mutual exclusivity, that is, excluding those items whose label they already known. As before, target words were represented by different pictures from those presented in the story. There were practice trials using familiar non-target words (e.g., dog, cat, spoon, pencil). For each target item, children received 1 point for a correct response. Chance level in this task was 25%.

Procedure

Dyads were invited to the University Infant and Child Study Centre. The study took place in an observation room that enabled non-intrusive audio and video recording of caregiver-child shared reading interactions. After greeting the participating dyad and establishing a rapport, the experimenter administered the first subtest of the visual attention assessment (NEPSY; Korkman et al., 1998) to the child, while the caregiver completed the vocabulary checklist of the target words. Subsequently, caregiver-child dyads were instructed to read the two books. Dyads received the following instructions: "*I would like you to read together as you would do at home. Please take your time, I will be back when you are finished*". The dyads sat on a sofa and were left alone in the observation room. The experimenter observed and video-recorded the shared reading interactions from the control room.

Each dyad read a book presented on paper and a different book presented on an iPad, with the order of presentation counterbalanced across participants. After the shared reading activity, the child was administered the second subtest of the visual attention assessment, the target vocabulary learning measures, and a picture

sequence task[1], while caregivers completed the BRIEF-P (Gioia et al., 2003). Children completed the standardised vocabulary assessment and the working memory assessments, either after a short break (*n* = 71) or within 12 days from the first visit (*n* = 29), depending on child compliance and dyad availability.

Analytic Plan

Statistical Models

We fit separate Generalized Linear Mixed Models (GLMM; (Baayen et al., 2008) for each word learning measure (i.e., phonological recall, semantic recall, and semantic recognition). We specified binomial error structure and logit link function for binary outcomes (i.e., phonological recall, semantic recognition), and Poisson family for count data (i.e., semantic recall). The full random effect structure supported by the data was included in the models (Matuschek et al., 2017). Convergence issues were addressed according to a stepwise procedure: first by increasing the number of iterations, then using different optimisers, and, as a last step, by simplifying the random effect structure.

Research Question 1. To investigate the effect of book format (print vs. digital) on word learning, we compared a model (M_1) including our test predictor (book format) with a model (M_0) lacking our test predictor but being otherwise identical. Both models controlled for prior vocabulary knowledge, executive function, age, and gender[2], all entered as fixed effects, to estimate any unique effect of book format, above and beyond other established predictors of word learning.

(M_0) Word Learning ~ Prior Vocabulary Knowledge + Executive function + Age + Gender + Random Effects

(M_1) Word Learning ~ *Book Format* + Prior Vocabulary Knowledge + Executive function + Age + Gender + Random Effects

Research Question 2. To investigate the interaction of book format and child characteristics (prior vocabulary knowledge and executive function) on word learning, we compared a model (M_2) including the interaction terms of interests with a model (M_1) lacking these interaction terms but being otherwise identical. Both models also

[1] The picture sequence task is analysed and discussed in separate manuscript.

[2] Here we use the term gender in line with the previous literature and to acknowledge the fact that language literacy development is embedded in a socio-cultural context. Strictly speaking however, we classed our child participants based on caregiver reported biological sex.

controlled for age and gender, specified as fixed effects, to conduct a stringent test of the hypothesized interactions.

(M_2) Word Learning ~ *Book Format*Prior Vocabulary Knowledge + Book Format*Executive function* + Age + Gender + Random Effects

Transformations

Our independent variable (IV) was categorical with two levels: print and digital. The IV was dummy-coded with print as the reference category. Age in months was z-transformed to facilitate the interpretation of model coefficients and to ease model convergence. There were moderate to high correlations between our measures of visual attention, forward digit span, backward digit span, and global executive function (see Table S2, Supplementary Materials). In this context and given the debated dimensionality of executive function in the pre-school years (Kapa & Erikson, 2020; Lerner & Lonigan, 2014; Montroy et al., 2019; Scionti & Marzocchi, 2021), we conducted a Principal Component Analysis (PCA) to assess the dimensionality of these measures. We adopted the procedure detailed by James et al (2021). Analyses were conducted in R with the package psych (Revelle, 2023). First, we checked the Kaiser-Meyer-Olkin measure of sampling adequacy. The overall Kaiser-Meyer-Olkin value for the data was 0.61, which is classified as mediocre (Kaiser, 1974). Then, we performed a PCA. A single executive function component emerged according to Wayne Velicer's Minimum Average Partial (MAP) criterion, and by retaining eigenvalues > 1. The proportion of variance explained by this component was 47%. As expected, loadings were positive for visual attention (.70), forward digit span (.74) and backward digit span (.80), and negative for global executive function (-.46), which provides an index of dysfunction, with higher scores indicating poorer executive skills. Given the extraction of a single component for executive function, a single score resulting from our four executive function measures was used in subsequent analyses.

Inference Criteria

To evaluate the significance of the contribution of our predictors, whilst avoiding multiple testing (Forstmeier & Schielzeth, 2011), we compared our full models with null models lacking our test predictors but being otherwise identical using the likelihood ratio test. The significance of the beta coefficients was indicated by $p <$ .05. Null hypothesis significance testing was complemented by an examination of odds

ratios (to provide an index of the strength of the relationships between our predictors and outcome) and relative confidence intervals in the case of logistic models. Marginal effects were plotted to provide straightforward visualization of predicted probabilities for the results of the more complex models (Lüdecke, 2018). Significant interaction terms accompanied by improvement in the model fit, as indicated by a significant likelihood ratio test, indicated the presence of conditional effects. In case of a discrepancy between the significance of the likelihood ratio test and that of model coefficients, the confidence intervals of model coefficients were inspected for symmetry and wideness to inform the interpretation. Marginal R^2 is reported to illustrate the proportion of variance explained by the fixed effects, while conditional R^2 is reported to illustrate the variance explained by both the fixed and random effects.

Implementation

The models were implemented in R version 4.1.3 (2022-03-10) with the function glmer of the R package lme4 (version 1.1- 33) (Bates et al., 2015). Predicted probabilities were computed using the function ggpredict of the R package ggeffects, version 1.3.2 (Lüdecke, 2018).

Transparency and Openness

Our research questions, hypotheses, design, sample size, and analytic plan were pre-registered prior to data collection on the Open Science Framework (OSF; https://osf.io/6uem9/?view_only=7bb996db9f7c4d54a7b3bf85723a0b16; https://osf.io/ancsx/?view_only=2dcb04be82564e40a3ce9f8d36856dc3). Any deviation from the pre-registration is detailed in this manuscript. Fully anonymised data and analyses scripts are available on the project's OSF repository.

Results

Descriptive statistics

Reading time across conditions was the same: On average, the print reading condition lasted 4 minutes and 33 seconds, and the digital reading condition 4 minutes and 34 seconds. On average, children correctly labelled 25% of the items in the phonological recall task; produced .87 units of meaningful information per item in the semantic recall task; and correctly identified 72% of the items in the semantic

recognition task (chance level of 25%). The means and standard deviations of our outcome measures grouped by condition (book format) are reported in Table 2.

Table 2 *Mean and Standard Deviations of Word Learning Outcomes by Book Format*

	Phonological recall[a]		*Semantic recall*[b]		*Semantic Recognition*[a]	
Book format	*M*	*SD*	*M*	*SD*	*M*	*SD*
Print	0.25	0.44	0.84	1.28	0.71	0.45
Digital	0.25	0.44	0.89	1.41	0.74	0.44

[a] binary outcome (0 incorrect, 1 correct. Range: 0 - 1)

[b] count outcome (number of meaningful units of information provided per target word. No upper limit; Range in the data: 0 - 7)

There were no obvious differences in phonological recall across book formats; in contrast, for semantic recall and semantic recognition, the mean accuracy in the digital condition was higher than that in the print condition. The means, standard deviations, and correlations between our quantitative predictors (collapsed over book format) and outcomes are reported in Table 3.

Table 3 *Means, Standard Deviations, and Correlations with Confidence Intervals*

Variable	*M*	*SD*	1	2	3	4	5
1. Age (months)	57.77	7.24					
2. Vocabulary (scaled score)	10.15	2.81	.03 [-.04, .10]				
3. Executive function (factor component)	0.00	1.00	.48** [.43, .53]	.22** [.15, .28]			
4. Phonological Recall[a]	0.25	0.44	.07* [.00, .14]	.10** [.03, .16]	.09* [.02, .15]		
5. Semantic Recall[b]	0.87	1.35	.09* [.02, .17]	.19** [.12, .27]	.12** [.04, .19]	.46** [.40, .51]	
6. Semantic Recognition[a]	0.72	0.45	.01 [-.06, .08]	.15** [.08, .22]	.08* [.02, .15]	.28** [.22, .34]	.23** [.16, .30]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

[a] binary outcome (0 incorrect, 1 correct. Range: 0 - 1)

[b] count outcome (number of meaningful units of information provided per target word. No upper limit; Range in the data: 0 - 7)

The correlations indicate significant, but weak, correlations between child characteristics and word learning measures. The means and standard deviations of our outcome measures grouped by gender are reported in Table S3, in Supplementary Materials. Accuracy by target word across word learning measures is depicted in figure S1, in the Supplementary Materials.

Preliminary analyses were conducted to check for the presence of any possible order effects on word learning, both in terms of book format (print or digital presented first) and story presentation (story A or B presented first). Order effects were not found for any of the word learning outcomes. There was no order effect on phonological recall (format: β = .02, *SE* = .16, *p* = .88; story: β = .12, *SE* = .16, *p* = .48), sematic recall (format: β = .07, *SE* = .19, *p* = .70; story: β = .24, *SE* = .19, *p* = .19), nor semantic recognition (format: β = .17, *SE* = .21, *p* = .41; story: β = - 0.07, *SE* = .21, *p* = .75).

Confirmatory Analyses

Phonological Recall

Model results are reported in Table 4.

Table 4 *Results of the Models Estimating the Effect of Book Format, Child Characteristics, and Their Interaction, on Phonological Recall*

	Phonological Recall											
	M_0				**M_1**				**M_2**			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-2.00	0.34	-2.68 – -1.33	**<0.001**	-2.01	0.35	-2.70 – -1.31	**<0.001**	-2.16	0.49	-3.11 – -1.21	**<0.001**
Vocabulary	0.07	0.03	0.01 – 0.14	**0.017**	0.07	0.03	0.01 – 0.14	**0.017**	0.09	0.04	0.00 – 0.18	**0.045**
Executive Function	0.14	0.10	-0.05 – 0.33	0.147	0.14	0.10	-0.05 – 0.33	0.147	0.14	0.13	-0.11 – 0.40	0.265
Age (months)	0.10	0.09	-0.08 – 0.28	0.297	0.10	0.09	-0.08 – 0.28	0.297	0.10	0.09	-0.08 – 0.28	0.298
Gender [M]	0.34	0.17	0.01 – 0.66	**0.045**	0.34	0.17	0.01 – 0.66	**0.045**	0.34	0.17	0.01 – 0.66	**0.045**
Book format [digital]					0.01	0.16	-0.31 – 0.33	0.949	0.31	0.66	-0.99 – 1.61	0.642
Book format [digital] * Vocabulary									-0.03	0.06	-0.15 – 0.09	0.645
Book format [digital] * Executive Function									-0.01	0.17	-0.34 – 0.33	0.975

Random Effects

σ^2	3.29	3.29	3.29
τ_{00}	0.00 $_{child\_ID}$	0.00 $_{child\_ID}$	0.00 $_{child\_ID}$
τ_{11}	0.00 $_{child\_ID.book\_format}$	0.00 $_{child\_ID.book\_format}$	0.00 $_{child\_ID.book\_format}$
ρ_{01}			
N	100 $_{child\_ID}$	100 $_{child\_ID}$	100 $_{child\_ID}$
Observations	807	807	807
Marginal R^2 / Conditional R^2	0.034 / NA	0.034 / NA	0.034 / NA

The likelihood ratio test comparing the model (M_1) with the test predictor of interest with the model (M_0) lacking the test predictor, indicated no significant main effect of book format on phonological recall (χ^2 = .004, *df* = 1, *p* = .95). Similarly, the comparison of the model (M_2) including the interaction between book format and child characteristics with the model (M_1) lacking these interaction terms, indicated no evidence of interactions between book format and child characteristics on phonological recall (χ^2 =.228, *df* = 2, *p* =.89). In sum, there was no evidence for any unique or mediated effect of book format on phonological recall.

There was evidence of a significant positive effect of prior vocabulary knowledge and gender on phonological recall across all models. Greater prior vocabulary knowledge significantly predicted phonological recall, as indicated by the positive sign of the respective coefficient, and boys were significantly more accurate than girls. Odds ratios are reported in Table S4, Supplementary Materials.

Semantic Recall

Model results are reported in Table 5.

Table 5 *Results of the Models Estimating the Effect of Book Format, Child Characteristics, and Their Interaction, on Semantic Recall*

	Semantic Recall											
	M_0				**M_1**				**M_2**			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-2.01	0.35	-2.68 – -1.33	**<0.001**	-1.99	0.35	-2.68 – -1.31	**<0.001**	-2.18	0.40	-2.97 – -1.38	**<0.001**
Vocabulary	0.12	0.03	0.06 – 0.18	**<0.001**	0.12	0.03	0.06 – 0.18	**<0.001**	0.14	0.04	0.07 – 0.21	**<0.001**
Executive Function	0.23	0.09	0.05 – 0.42	**0.014**	0.23	0.09	0.05 – 0.41	**0.015**	0.18	0.11	-0.03 – 0.39	0.097
Age (months)	0.07	0.09	-0.10 – 0.24	0.435	0.07	0.09	-0.10 – 0.24	0.439	0.07	0.09	-0.10 – 0.24	0.396
Gender [M]	0.47	0.16	0.16 – 0.78	**0.003**	0.47	0.16	0.16 – 0.78	**0.003**	0.50	0.16	0.18 – 0.81	**0.002**
Book format [digital]					-0.02	0.15	-0.31 – 0.26	0.868	0.40	0.39	-0.36 – 1.15	0.307
Book format [digital] * Vocabulary									-0.03	0.03	-0.10 – 0.03	0.331
Book format [digital] * Executive Function									0.11	0.09	-0.08 – 0.29	0.246

Random Effects

σ^2	0.98	0.98	0.94
τ_{00}	0.33 $_{child\_ID}$	0.41 $_{child\_ID}$	0.35 $_{child\_ID}$
τ_{11}	0.64 $_{child\_ID.book\_format}$	0.65 $_{child\_ID.book\_format}$	
ρ_{01}	0.17 $_{child\_ID}$	-0.46 $_{child\_ID}$	
ICC	0.25	0.34	0.27
N	95 $_{child\_ID}$	95 $_{child\_ID}$	95 $_{child\_ID}$
Observations	676	676	676
Marginal R^2 / Conditional R^2	0.151 / 0.365	0.136 / 0.426	0.163 / 0.392

Similar to phonological recall, the likelihood ratio test comparing the model (M_1) with the model (M_0) indicated that there was no significant main effect of book format on semantic recall (χ^2 = .028, *df* = 1, *p* = .87), and the likelihood ratio test comparing the model (M_2) with the model (M_1) did not support the presence of significant interactions between book format and child characteristics on semantic recall (χ^2 =2.10, *df* = 2, *p* =.35).

In line with the results reported for phonological recall, there was a significant positive main effect of prior vocabulary knowledge and gender on semantic recall across models (Table 6). Greater prior vocabulary knowledge significantly predicted semantic recall, as indicated by the positive sign of the respective coefficient, and boys were significantly more accurate than girls. Different to the results for phonological recall, there was evidence for a significant main effect of executive function on semantic recall in both M_1 and M_2 models.

Semantic recognition

The effect of book format on semantic recognition was examined using the same model fitting process and comparison as before. Model results are reported in Table 6.

Table 6 *Results of the Models Estimating the Effect of Book Format, Child Characteristics, and Their Interaction, on Semantic Recognition*

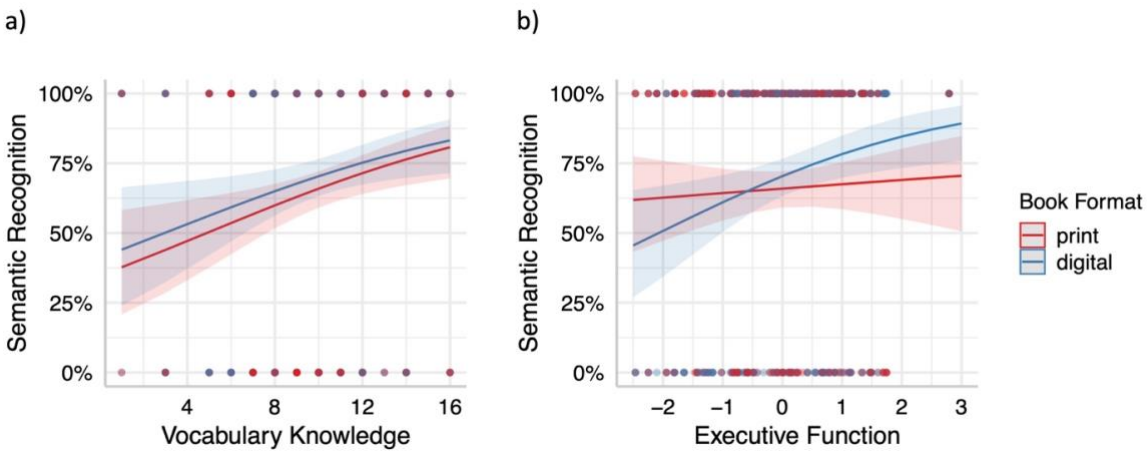
	Semantic Recognition											
	M_0				**M_1**				**M_2**			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-0.48	0.38	1.23 – 0.26	0.200	-0.57	0.39	1.33 – 0.18	0.138	-0.63	0.47	1.55 – 0.29	0.178
Vocabulary	0.12	0.04	0.06 – 0.19	**<0.001**	0.13	0.04	0.06 – 0.19	**<0.001**	0.13	0.04	0.04 – 0.22	**0.004**
Executive Function	0.23	0.12	0.00 – 0.47	0.053	0.22	0.12	0.01 – 0.46	0.063	0.07	0.14	0.20 – 0.34	0.609
Age (months)	-0.10	0.11	0.32 – 0.12	0.365	-0.10	0.11	0.32 – 0.12	0.372	-0.10	0.11	0.32 – 0.12	0.372
Gender [M]	0.65	0.21	0.25 – 1.06	**0.002**	0.66	0.21	0.25 – 1.06	**0.001**	0.66	0.21	0.26 – 1.06	**0.001**
Book format [digital]					0.17	0.19	0.20 – 0.54	0.362	0.27	0.64	0.98 – 1.51	0.672
Book format [digital] * Vocabulary									-0.01	0.06	0.13 – 0.12	0.920
Book format [digital] * Executive Function									0.35	0.18	0.00 – 0.69	**0.047**
Random Effects												
σ^2	3.29				3.29				3.29			
τ_{00}	0.22 $_{child\_ID}$				0.22 $_{child\_ID}$				0.22 $_{child\_ID}$			
τ_{11}	0.28 $_{child\_ID.book\_format}$				0.26 $_{child\_ID.book\_format}$				0.16 $_{child\_ID.book\_format}$			

ρ_{01}	0.25 $_{child\_ID}$	0.51 $_{child\_ID}$	0.75 $_{child\_ID}$
ICC	0.06	0.06	0.06
N	100 $_{child\_ID}$	100 $_{child\_ID}$	100 $_{child\_ID}$
Observations	802	802	802
Marginal R^2 / Conditional R^2	0.068 / 0.126	0.070 / 0.128	0.080 / 0.138

The likelihood ratio test indicated no significant main effect of book format on semantic recognition (χ^2 = .828, *df* = 1, *p* = .36), and there was no evidence of interactions between book format and child characteristics on semantic recognition (χ^2 = 3.99 *df* = 2, *p* =.14). However, the results of the more complex model (M_2) supported the presence of a significant interaction term between book format and executive function as illustrated by the model coefficient. Further, R^2 was larger in M_2 compared to M_1 (Table 6).

To inform our interpretation, we examined whether the confidence interval of the significant interaction term was too wide or asymmetrical, which would suggest giving more weight to the results of the likelihood ratio test (Field et al., 2012, p.72). As this was not the case, we followed a data-driven approach and pruned the nonsignificant interaction between book format and vocabulary in a follow-up model (M_3). We then contrasted the model (M_3) containing only the interaction between book format and executive function against a model (M_1) lacking this term but being otherwise identical. The results of the likelihood ratio test supported the presence of a significant interaction between book format and executive function (χ^2 =3.98, *df* = 1, *p* =.05). Further, the coefficients of M_3 confirmed the presence of a significant interaction between book format and executive function as well as main effects of vocabulary and gender (see output for M_3 in Table S6, Supplementary Materials). In sum, there was a strong indication that book format interacted with executive function, but not with vocabulary knowledge, for semantic recognition. To guide interpretation of the interaction, predicted probability plots were generated (Fig. 2).

Figure 2 *Predicted Values of Semantic Recognition by a) Vocabulary Knowledge and b) Executive Function, Across Book Formats*



For children with higher executive function skills, digital-based shared reading had a positive effect on semantic recognition compared to print-based shared reading, while the opposite was true for children with lower executive function. Looking at this interaction from another perspective, the effect of executive function on semantic recognition was stronger in the digital condition, compared to the print condition, while prior vocabulary was a strong determinant of semantic recognition across both book formats.

We note that there was a significant positive main effect of prior vocabulary knowledge and gender on semantic recognition across models. Greater prior vocabulary knowledge significantly predicted semantic recognition, as indicated by the positive sign of the respective coefficient. Further, boys were significantly more accurate than girls in the semantic recognition task. Odds ratios are reported in Table S5, in the Supplementary Materials.

Discussion

Our primary aim was to investigate evidence for any unique and mediated effects of book format (print vs. digital) on word learning from shared book reading in 4- to 5-year-old children. This study advances our understanding of word learning from naturalistic shared reading in several ways. Our results suggest that book format influenced only certain aspects of word learning and that this was conditional on child characteristics. Specifically, digital-based shared reading led to better semantic recognition in children with higher executive function, and lower semantic recognition in children with lower executive function. Further, children with greater prior vocabulary knowledge and boys were significantly more accurate across word learning measures. Lastly, executive function uniquely contributed to semantic recall. We discuss each of these findings as well as their implications for theory and practice.

Similar to the meta-analysis by Furenes et al. (2021) we did not find evidence of a main effect of book format on word learning from shared reading in young children. However, for semantic recognition, there was evidence that book format interacted with executive function. One could interpret these findings in relation to the cognitive theory of multimedia learning, which emphasises the importance of reducing cognitive load to optimise learning (Mayer, 2003; Mayer & Moreno, 2003). In the context of the present study, handling and operating the digital device, albeit with caregiver support, may have led to an increase in cognitive load, which was particularly detrimental for children with immature executive function (Courage, 2019). At the physiological level, adults reading on screen compared to print present with overactivity in the pre-frontal cortex, a brain region strongly implicated in executive functions (Honma et al., 2022). This supports the interpretation that the digital format may have posed greater demands on executive functions.

Children with higher executive function benefitted from the digital format, suggesting that once executive function is sufficiently developed, the digital format may bring some advantages. A possible explanation for such an advantage can be the relative novelty of digital-based shared reading for our sample (see Table S1, Supplementary Materials). This may have boosted children's interest in the activity and, in turn, learning (Flewitt et al., 2015; O'Toole & Kannass, 2018). Prior research

indeed suggests that children who rarely use a tablet, learn more from the digital format, but the opposite is true for weekly tablet users (Reich et al., 2019).

There was a significant and positive main effect of prior vocabulary knowledge on word learning, in line with a wealth of prior research (Sénéchal et al., 1995). Thus, regardless of format, prior vocabulary is an important determinant of word learning from shared reading. These findings relate to the so-called Matthew Effects, referring to rich-get-richer and poor-get-poorer processes (Stanovich, 1986). Children with larger vocabulary are better equipped to understand the story and therefore have more cognitive resources available to process new words as they encounter them. Further, their existing knowledge base and category knowledge can facilitate the acquisition of new words in toddlers (Borovsky et al., 2016).

In line with accumulating evidence (Hadley et al., 2021; Kapa & Erikson, 2020), we found that executive function predicted semantic recall after controlling for prior vocabulary knowledge. This further confirms that higher-level cognitive functions are particularly important in the acquisition of word meanings as assessed with a definition task. From a theoretical perspective, this is in line with theories that emphasise the crucial role of broader social-cognitive skills in word learning and language acquisition (Tomasello, 2003), and also with models of reading comprehension that recognise the role of higher-level cognitive skills and processing resources (Cain, Oakhill, et al., 2004; Y.-S. G. Kim, 2020).

An unexpected finding was that boys outperformed girls on all word learning measures. This gender effect was small to medium sized (with OR ranging from 1.4 to 1.98), but reliable, across word learning measures. Girls are typically reported to outperform boys in language and literacy development measures (Eriksson et al., 2012; Frank et al., 2017; S. Logan & Johnston, 2010), an effect that is small, but reliably found across cultural contexts (Eriksson et al., 2012). Some researchers have pointed to the importance of understanding in which specific contexts gender differences are observed (Bergman Deitcher et al., 2019; S. Logan & Johnston, 2010). For instance, it has been reported that boys outperform girls in word learning from shared reading of informational, but not narrative, texts (Bergman Deitcher et al., 2019). Boys in our sample were not significantly older, nor had higher verbal ability or executive function.

Therefore, characteristics of our setting and task may explain the advantage of boys in word learning, an area of inquiry that deserves further investigation.

Implications and practical significance

From a theoretical perspective, these findings align with the capacity model (Fisch, 2000) and the differential susceptibility to media effect model (Valkenburg & Peter, 2013), which highlight the need to take individual differences into account when studying learning from different media. There are several strengths of our study design, as noted in the Methods, to endorse these findings as robust and reliable. These include: our use of a within-subjects design and large sample to enhance statistical power relative to previous research; our use of the same reader to narrate the text; three outcome measures to assess both phonological and semantic features of word learning, and a semi-naturalistic paradigm to support generalisation to real-world settings. Of course, these findings still require replication. If found, one obvious conclusion is that the use of print-based reading materials should be favoured amongst children with still immature executive functioning, and potentially also in developmental conditions characterised by poor executive function.

Limitations and future directions

This study comes with limitations and suggestions for future research, in addition to those already noted. First, we note that only a relatively small sample of books and word types was included in this study's materials. Whilst this is a common feature in the shared reading literature, it is important that future studies incorporate different books and word types to increase the generalisability of findings beyond specific sets of stimuli. In this context, sharing e-books under a Creative Commons license, as we have done, may help researchers working under different time constraints to efficiently adapt existing materials to their research needs, and thus include larger sets of stimuli in future studies.

Another limitation concerns the relatively short time scale across which learning was assessed. Learning novel words and their meanings happens across multiple exposures over time, which should be the aim of future work. These findings are also limited to the cultural context of our participating sample, which was recruited within travelling distance of our child assessment labs. Further investigation is needed to generalise these findings to different cultural contexts, especially considering that

these results may vary as a function of experience with digital media. Age and gender differences were not the focus of this study but clearly deserve further investigation given the findings reported here. Finally, a different, yet complementary area of research, concerns the investigation of the affordances of digital books that open up new opportunities to scaffold understanding and learning. For instance, it is critical to investigate how e-book features can be designed to support parental mediation during shared reading (Troseth et al., 2020). This might be usefully informed by research on the role of built-in scaffolds to support understanding and learning during independent reading in beginning readers (Diprossimo et al., 2023).

Conclusions

Shared book reading provides unique opportunities to boost children's vocabulary knowledge, which lays the foundation for later educational and social outcomes. Our findings provide the first empirical evidence for an interplay between book format and executive function, suggesting that the effect of book format is not the same for all children. Considering the changing landscape of children's early literacy experiences, it is critical to build the evidence base to inform recommendations and guidelines around early literacy practices. This study adds to the evidence base by showing that when we consider different measures of word learning, we observe book format-related differences in learning that are contingent on child characteristics. These findings underscore the importance of considering multiple measures of word learning as well as individual differences both in future studies, and when developing guidelines for users of research.

Acknowledgements

We sincerely thank all the participating caregivers and children. We are grateful to Simya Aravamuthan, Gracey Caller, and Arwen Hon for their support with data scoring, inputting, and/or recruitment of participants. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Actions grant agreement No 857897.

CRediT Author Statement

Laura Diprossimo**:** Conceptualization, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. Kate Cain**:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

References

Adlof, S. M., Baron, L. S., Bell, B. A., & Scoggins, J. (2021). Spoken Word Learning in Children With Developmental Language Disorder or Dyslexia. *Journal of Speech, Language, and Hearing Research*, *64*, 2734–2749. https://doi.org/10.23641/asha

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, *105*(1), 158–173.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1 SE-Articles), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergman Deitcher, D., Johnson, H., & Aram, D. (2019). Does book genre matter? Boys' and girls' word learning from narrative and informational books in the preschool years. *Journal of Research in Reading*, *42*(1), 193–211. https://doi.org/10.1111/1467-9817.12266

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology*, *150*, 404–410. https://doi.org/10.1016/j.jecp.2016.06.009

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared Book Reading: When and How Questions Affect Young Children's Word Learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*, *19*(6), 918–932. https://doi.org/10.1111/desc.12343

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), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *4*, 284–290.

Courage, M. L. (2019). *From Print to Digital: The Medium Is Only Part of the Message* (pp. 23–43). https://doi.org/10.1007/978-3-030-20077-0_3

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, *25*, 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

Diprossimo, L., Ushakova, A., Zoski, J., Gamble, H., Irey, R., & Cain, K. (2023). The associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment: Insights from a big data approach. *Contemporary Educational Psychology*, *73*. https://doi.org/10.1016/j.cedpsych.2023.102165

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis. *Child Development*, *91*(2), e383–e399. https://doi.org/10.1111/cdev.13225

Dunn, L. M., & Dunn, D. M. (2009). The British picture vocabulary scale. In *GL Assessment Limited.*

Elley, W. B. (1989). Vocabulary Acquisition from Listening to Stories. *Reading Research Quarterly*, *24*(2), 174–187.

Eng, C. M., Tomasic, A. S., & Thiessen, E. D. (2019). Contingent Responsivity in E-Books Modeled From Quality Adult-Child Interactions: Effects on Children's Learning and Attention. *Developmental Psychology*, *56*(2), 285–297. https://doi.org/10.1037/dev0000869

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Gayraud, F., Kovacevic, M., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology*, *30*(2), 326–343. https://doi.org/10.1111/j.2044-835X.2011.02042.x

Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, *32*(3), 343–364. https://doi.org/10.1177/0142723711422626

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). The MacArthur Communicative Development Inventories: User's guide and technical manual. In *San Diego, CA: Singular Publishing Group.*

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R* (Sage publications).

Fisch, S. M. (2000). A capacity model of children's comprehension of educational content on television. *Media psychology*, *2*(1), 63-91. https://doi.org/10.1207/S1532785XMEP0201_4

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The Effects of Shared Storybook Reading on Word Learning: A Meta-Analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/http://dx.doi.org/10.1037/dev0000512

Flewitt, R., Messer, D., & Kucirkova, N. (2015). New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy*, *15*(3), 289–310. https://doi.org/10.1177/1468798414533560

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A Comparison of Children's Reading on Paper Versus Screen: A Meta-Analysis. *Review of Educational Research*, *91*(4), 483–517. https://doi.org/10.3102/0034654321998074

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological Short-Term Memory and New Word Learning in Children. *Developmental Psychology*, *33*(6), 966–979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological Short-term Memory and Vocabulary Development: Further Evidence on the Nature of the Relationship. *Applied Cognitive Psychology*, *13*(1), 65–77.

https://doi.org/10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O

Gioia, G. A. , Espy, K. A. , & Isquith, P. K. (2003). BRIEF-P: Behavior rating inventory of executive function - preschool version. In *Psychological Assessment Resources (PAR).*

Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Kim, E., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Journal of Applied Developmental Psychology*, *77*, 101332. https://doi.org/10.1016/j.appdev.2021.101332

Hadley, E. B., & Dickinson, D. K. (2020). Measuring young children's word knowledge: A conceptual review. *Journal of Early Childhood Literacy*, *20*(2), 223–251. https://doi.org/10.1177/1468798417753713

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181–198. https://doi.org/10.1002/rrq.130

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*, 104265. https://doi.org/10.1016/j.cognition.2020.104265

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4–14. https://doi.org/10.1037/a0027238

Honma, M., Masaoka, Y., Iizuka, N., Wada, S., Kamimura, S., Yoshikawa, A., Moriya, R., Kamijo, S., & Izumizaki, M. (2022). Reading on a smartphone affects sigh generation, brain activity, and comprehension. *Scientific Reports*, *12*(1), 1–8. https://doi.org/10.1038/s41598-022-05605-0

James, E., Currie, N. K., Tong, S. X., & Cain, K. (2021). The relations between morphological awareness and reading comprehension in beginner readers to young adolescents. *Journal of Research in Reading*, *44*(1), 110–130. https://doi.org/10.1111/1467-9817.12316

Kaiser, H. F. (1974). An index of factorial simplicity. *psychometrika*, *39*(1), 31-36.

Kapa, L. L., & Erikson, J. A. (2020). The relationship between word learning and executive function in preschoolers with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *63*(7), 2293–2307. https://doi.org/10.1044/2020_JSLHR-19-00342

Kim, Y.-S. G. (2020). Hierarchical and Dynamic Relations of Language and Cognitive Skills to Reading Comprehension: Testing the Direct and Indirect Effects Model of Reading (DIER). *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000407.supp

Korat, O., & Or, T. (2010). How new technology influences parent-child interaction: The case of e-book reading. *First Language*, *30*(2), 139–154. https://doi.org/10.1177/0142723709359242

Korkman, M., Kirk, U., & Kemp, S. (1998). NESPY: A developmental neuropsychological assessment. In *San Antonio, TX: The Psychological Corporation.*

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*(4), 978–990.

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes. *Journal of Speech, Language, and Hearing Research*, *52*, 1401–1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2020). More than words: Narrator engagement during storytelling increases children's word learning, story comprehension, and on-task behavior. *Early Childhood Research Quarterly*, *51*, 338–351. https://doi.org/10.1016/j.ecresq.2019.12.009

Lerner, M. D., & Lonigan, C. J. (2014). Executive Function Among Preschool Children: Unitary Versus Distinct Abilities. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 626–639. https://doi.org/10.1007/s10862-014-9424-3

Logan, J. A. R., Justice, L. M., Yumus¸ Melike, & Chaparro-Moreno, L. J. (2019). When Children Are Not Read to at Home: The Million Word Gap. *Journal of Developmental Behavioral Pediatrics*, *40*(5), 383–386. DOI: 10.1097/DBP.0000000000000657

Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, *62*(2), 175–187. https://doi.org/10.1080/00131911003637006

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/s0959-4752(02)00016-6

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, *39*(5), 527–546. https://doi.org/10.1177/0142723719849996

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Montroy, J. J., Merz, E. C., Williams, J. M., Landry, S. H., Johnson, U. Y., Zucker, T. A., Assel, M., Taylor, H. B., Lonigan, C. J., Phillips, B. M., Clancy-Menchetti, J., Barnes, M. A., Eisenberg, N., Spinrad, T., Valiente, C., de Villiers, J., & de Villiers, P. (2019). Hot and cool dimensionality of executive function: Model invariance across age and maternal education in preschool children. *Early Childhood Research Quarterly*, *49*, 188–201. https://doi.org/10.1016/j.ecresq.2019.06.011

Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. (2019). Parent-Toddler Social Reciprocity during Reading from Electronic Tablets vs Print Books. *JAMA Pediatrics*, *173*(11), 1076–1083. https://doi.org/10.1001/jamapediatrics.2019.3480

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*(October 2018), 100290. https://doi.org/10.1016/j.edurev.2019.100290

O'Toole, K. J., & Kannass, K. N. (2018). Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology*, *173*, 100–115. https://doi.org/10.1016/j.jecp.2018.03.013

Ozturk, G., & Hill, S. (2020). Mother–child interactions during shared reading with digital and print books. *Early Child Development and Care*, *190*(9), 1425–1440. https://doi.org/10.1080/03004430.2018.1538977

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*, 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001

Price, J., & Kalil, A. (2019). The Effect of Mother–Child Reading Time on Children's Reading Skills: Evidence From Natural Within-Family Variation. *Child Development*, *90*(6), e688–e702. https://doi.org/10.1111/cdev.13137

Reich, S. M., Yau, J. C., Xu, Y., Muskat, T., Uvalle, J., & Cannata, D. (2019). Digital or Print? A Comparison of Preschoolers' Comprehension, Vocabulary, and Engagement From a Print Book and an e-Book. *AERA Open*, *5*(3), 233285841987838. https://doi.org/10.1177/2332858419878389

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. *Northwestern University, Evanston, Illinois.*

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The Pace of Vocabulary Growth Helps Predict Later Vocabulary Skill. *Child Development*, *83*(2), 508–525. https://doi.org/10.1111/j.1467-8624.2011.01710.x

Salmerón, L., Altamura, L., Delgado, P., Karagiorgi, A., & Vargas, C. (2023). Reading comprehension on handheld devices versus on paper: A narrative review and meta-analysis of the medium effect and its moderators. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000830

Salmerón, L., Delgado, P., Vargas, C., & Gil, L. (2021). Tablets for all? Testing the screen inferiority effect with upper primary school students. *Learning and Individual Differences*, *86*. https://doi.org/10.1016/j.lindif.2021.101975

Sarı, B., Başal, H. A., Takacs, Z. K., & Bus, A. G. (2019). A randomized controlled trial to test efficacy of digital enhancements of storybooks in support of narrative comprehension and word learning. *Journal of Experimental Child Psychology*, *179*, 212–226. https://doi.org/10.1016/j.jecp.2018.11.006

Scionti, N., & Marzocchi, G. M. (2021). The dimensionality of early executive functions in young preschoolers: Comparing unidimensional versus bidimensional models and their ecological validity. *Child Neuropsychology*, *27*(4), 491–515. https://doi.org/10.1080/09297049.2020.1868419

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual Differences in 4-Year-Old Children's Acquisition of Vocabulary During Storybook Reading. *Journal of Educational Psychology*, *87*(2), 218–229. https://doi.org/10.1037/0022-0663.87.2.218

Shi, J., Gu, Y., & Vigliocco, G. (2022). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*. https://doi.org/10.1111/desc.13357

Stanovich, K. E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/rrq.21.4.1

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and Pitfalls of Multimedia and Interactive Features in Technology-Enhanced Storybooks: A Meta-Analysis. *Review of Educational Research*, *85*(4), 698–739. https://doi.org/10.3102/0034654314566989

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. In *Harvard University Press*.

Troseth, G. L., Strouse, G. A., Flores, I., Stuckelman, Z. D., & Russo Johnson, C. (2020). An enhanced eBook facilitates parent–child talk during shared reading by families of low socioeconomic status. *Early Childhood Research Quarterly*, *50*, 45–58. https://doi.org/10.1016/j.ecresq.2019.02.009

Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221-243. https://doi.org/10.1111/jcom.12024

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., & et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559. https://doi.org/10.1037//0012-1649.24.4.552

Wiig, E. H. , Secord, W. A. , & Semel, E. (2004). Clinical evaluation of language fundamentals—Preschool, (CELF Preschool-2). In *Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.*

Chapter 3: Caregivers' Fine-Tuning and its Effects on Children's Word Learning During Shared Book Reading

Linking Statement

In this pre-registered study (OSF Project https://osf.io/g9dw8/; Registration https://doi.org/10.17605/OSF.IO/4CM73), an in-depth analysis of caregivers' scaffolding was conducted. The rich corpus of video recordings of caregiver-child shared reading interactions collected in the context of the study reported in Chapter 2 was used. This enabled the examination of book-format-related differences in caregivers' scaffolding. The generalisability of the fine-tuning hypothesis was examined, that is, caregivers' ability to tailor scaffolds to their child's individual lexical knowledge to support word learning. This study is a natural complement to the study reported in Chapter 2, as it explores in detail the contribution of caregivers' scaffolding to children's word learning, and also how scaffolding is affected by different reading media.

Author Contribution Statement

In the chapter entitled “Caregivers’ Fine-Tuning and its Effects on Children’s Word Learning During Shared Book Reading” the authors agree to the following contributions:

Laura Diprossimo, 80% (Conceptualization, Formal analysis, Data curation, Writing – original draft, Visualization, Project administration)

Kate Cain, 20% (Conceptualization, Writing – original draft, Supervision, Funding acquisition)

Abstract

This study tested the generalisability of the fine-tuning hypothesis to word learning in the context of shared reading. Caregivers' behaviour in two storybook presentation formats (print, digital) was also compared. English-speaking dyads were videotaped during shared reading (N_{dyads}= 78, mostly White; Caregiver M_{age} = 37.69 years, 74 women; Child M_{age} = 57.74 months, 43 girls; Data collected in the UK in 2023). Before shared reading, caregivers completed a vocabulary checklist to assess their children's knowledge of the target words embedded in the storybooks. Word learning was measured after shared reading. Caregivers adjusted verbal, but not gestural, scaffolds to their children's reported lexical knowledge. Fine-tuning predicted semantic recall. Overall, caregivers provided significantly fewer scaffolds with digital compared to print books.

Keywords: word learning, fine-tuning, child-directed language, scaffolding, gesture, digital media, shared book reading.

Introduction

Children learn new words at an incredible rate. Their vocabulary acquisition is enhanced by the rich multimodal cues provided by caregivers in play and conversation, and extended through encounters with written language during shared reading (Hadley & Dickinson, 2019; Masek et al., 2021; Montag et al., 2015; Tomasello, 2003). In recent years, there has been an increasing interest in understanding the dynamic and interactive mechanisms that support children's word learning and language development (Donnellan et al., 2023; Fusaroli et al., 2023; Leung et al., 2021; Shi et al., 2022). This recent work highlights the importance of understanding the ecological niche in which language development unfolds, and suggests that studying young children as isolated learners will provide, at best, an incomplete account of the mechanisms that support their language learning.

Caregivers consistently adapt their language when talking to children compared to adults (Fernald & Mazzie, 1991; Hills, 2013; Newport & Gleitman, 1984; Onnis et al., 2008). Both the prosodic and syntactic properties of child-directed language have been shown to support word learning in young children (Graf Estes & Hurley, 2013; Ma et al., 2011; Schwab & Lew-Williams, 2016). Crucially, the properties of child-directed language change as children age (Huttenlocher et al., 2010). This adaptive and flexible nature of caregivers' input is proposed to support children's language learning, by providing them with input at the appropriate level of complexity (Snow, 1972), a view which is in line with the *scaffolding* principle (Wood et al., 1976) and the notion of the *zone of proximal development* (Vygotsky, 1978). What is not known is whether caregivers adapt their communication to their children's general linguistic knowledge, or modulate it at a more fine-grained level, as a function of their children's knowledge of specific words. The focus of this study was to examine caregivers' communicative modulations during shared reading with their 4- to 5-year-olds, and its influence on children's learning of the meanings of unknown words.

Fine-Tuning and Word Learning

The fine-tuning hypothesis proposes that caregivers adapt their communication to their children's individual lexical knowledge in a way that supports their word learning (Leung et al., 2021). For instance, Leung et al. (2021) showed that caregivers adjust the length and content of their referring expressions to their two-year-olds'

individual lexical knowledge during a tablet-based reference game. Specifically, caregivers provided longer referring expressions for words they thought their children did not know. Longer referring expressions predicted children's accuracy only for words that caregivers thought their children did not know, suggesting that caregivers' fine-tuning contributed to children's success in the reference game. The content of referring expressions also differed by word knowledge: Caregivers provided more comparisons, descriptors, and superordinate category labels for unknown compared to known words, while the opposite was true for subordinate category labels. In a study of toy-play interactions, Donnellan et al. (2023) found that caregivers' questions asking and pointing gestures were also adapted to 3- to 4-year-olds' individual lexical knowledge and supported vocabulary acquisition. Specifically, asking more questions for unknown compared to known words supported immediate word learning, while pointing more towards known than unknown words supported immediate word learning.

There is therefore emerging evidence suggesting that caregivers fine-tune their verbal and gestural behaviours to their child's individual lexical knowledge to support word learning (Donnellan et al., 2023; Leung et al., 2021). To the best of our knowledge, only two studies have investigated caregivers' verbal and gestural fine-tuning and its effects on children's word learning. Our study focuses on another semi-naturalistic interaction that supports word learning. Whether these recent findings from other contexts would readily generalise to shared reading remains unexplored and deserves further investigation.

Shared Reading and Scaffolding Word Learning

Shared reading is an important context for word learning for several reasons. First, the language of storybooks is lexically richer than conversation and therefore provides learning opportunities beyond child-directed language (Montag et al., 2015). Second, adults' mediation supports word learning from shared reading (for a meta-analysis see Flack et al., 2018). Repetitions, definitions, and questions in extra-textual talk have been shown to support 3-to 6-year-olds' word learning from shared book reading (Blewitt et al., 2009; Flack et al., 2018; Lenhart et al., 2019). In line with embodied cognition accounts of word learning (e.g., Sadoski & Lawrence, 2023; Smith & Gasser, 2005), gestures have also been shown to support learning (Barnes et al.,

2023; Flack et al., 2018; Flack & Horst, 2018). Deictic gestures such as pointing facilitate 3.5-year-olds' word learning from shared reading in experimental contexts (Flack & Horst, 2018). During naturalistic shared reading in the classroom, meaning-focused gestures such as iconic and representational gestures, are associated with 4.5-year-olds' vocabulary development (Barnes et al., 2023).

In addition, scaffolding-like procedures, such as asking questions of gradually increasing complexity to the child, have been shown to facilitate 2- to 4-year-olds' acquisition of elaborated word meanings from shared reading (Blewitt et al., 2009). This underscores the importance of tailoring support to young children's lexical knowledge and suggests that fine-tuning may be an important mechanism for word learning in the shared reading context. Yet studies exploring whether and how caregivers adapt their support to children's lexical knowledge during shared reading are currently lacking.

Book Presentation Format

Contemporary reading experiences involve new media of story presentation, such as iPad and tablet, in addition to traditional print books (Kucirkova, 2019). Prior research suggests that the book presentation format (print vs digital) influences the quality of interactions during shared reading with 2- to 6-year-olds (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020). For instance, digital books are associated with less communicative initiations, responses and expanding talk by mothers (Korat & Or, 2010), and with lower social reciprocity between the parent-toddler dyads (Munzer et al., 2019). Yet it should be noted that digital books have been reported to enhance communicative initiations, responsiveness, and visual attention in children (Korat & Or, 2010; Richter & Courage, 2017; Wainwright et al., 2020). Against this background, we explored whether book format influenced caregivers' scaffolding during shared reading.

The Present Study

This study investigates scaffolding by English-speaking caregivers during shared reading interactions with their 4- to 5-year-olds. Shared reading frequency has been shown to be an important predictor of vocabulary knowledge and growth in this age group (Mol & Bus, 2011; Sénéchal et al., 2008). Specifically, we examined whether, to what extent, and how caregivers modulated their scaffolds as a function of their child's

individual lexical knowledge. We then assessed whether this modulation or *fine-tuning*, supported children's word learning. We used data from a corpus of videotaped shared reading interactions and coded for caregivers' provision of verbal and gestural scaffolds. Our dataset included responses to a vocabulary checklist completed by each caregiver before the shared reading interaction, which provided a proxy of their child's individual lexical knowledge. Measures of each child's target word learning and concurrent vocabulary knowledge were collected after the interaction.

If the fine-tuning hypothesis generalises to semi-naturalistic shared reading interactions, caregivers should display fine-tuning behaviours during shared reading, that is, they should modulate their scaffolds as a function of their children's lexical knowledge. Caregivers should provide more scaffolds for words that are unknown to their children, although this may vary across types of scaffolds (Donnellan et al., 2023). We expected to find evidence of fine-tuning across verbal and gestural scaffolds (Donnellan et al., 2023; Leung et al., 2021). If the digital reading medium negatively affects dyads' social reciprocity (Munzer et al., 2019), and hinders caregivers' communicative initiations, responsiveness, and expansions (Korat & Or, 2010), we should observe less scaffolding with digital storybooks, compared to print. Finally, fine-tuning should enhance children's word learning by providing input at the appropriate level of difficulty.

Methods

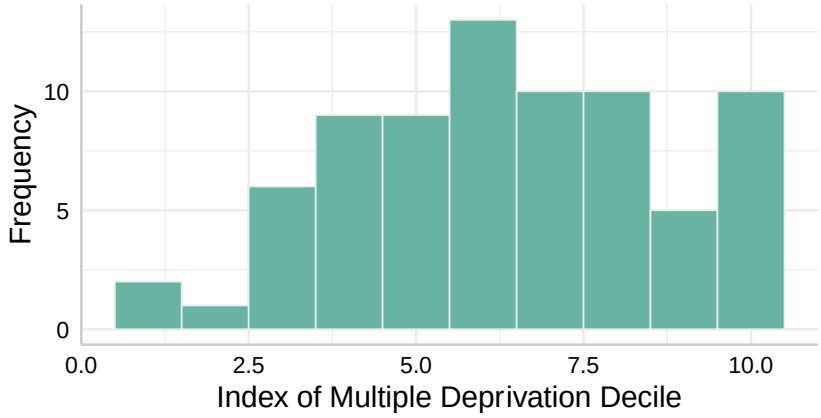
Participants

Seventy-eight British English-speaking caregiver-child dyads provided eligible data for this study. Data was collected between January and June 2023 in the context of a larger study (anonymised for review) which included 100 dyads. Overall, a small proportion of data relevant to the research aims of this paper was missing (5.8 %), and largely due to difficulties in performing the behavioural coding of specific video recordings, either due to equipment failure (the video recording was corrupted), or inaudible speech (see Figure S2 and S3 in the Supplementary Materials for the patterns of missingness). Previous studies addressing similar research questions with analogous analytic approaches have included 38 to 41 dyads (Leung et al., 2021; Shi et al., 2022). Therefore, we are confident that our sample provides sufficient power to detect the

effects under investigation, as stated in our pre-registration (https://osf.io/4cm73/?view_only=37ff877cb13a491bb424e2b07aebb0a5).

Children were typically developing, as reported by their caregivers, and aged between 48 to 71 months (M_{age} = 57.74; SD = 7.00; 43 girls). Caregivers were aged 29 to 45 years (M_{age} = 37.69; SD = 3.67; 74 self-reported as women; 4 as men) and were predominantly highly educated, with 63 achieving an undergraduate degree or higher. The socio-economic status (SES) of our participants was derived through their postcode (Government of the UK, 2019). We obtained a composite score, The Index of Multiple Deprivation, reflecting seven domains of deprivation, which include income, employment, health, education, barriers to housing and services, crime, and living environment (Ministry of Housing Communities and Local Government, 2019). We report the distribution of our sample according to the Index of Multiple Deprivation Decile in Figure 3.

Figure 3 *Index of Multiple Deprivation Decile Indexing SES of Participants*



The vast majority of dyads were White, reflecting the ethnic composition of the North West of England (85.6% White, 8.4 % Asian, 2.3% Black, 2.2% Mixed, 1.5% Other; Government of the UK, 2022). Participants were recruited via the university database, which covers a broad demographic, social media advertisements, and flyers distributed in public book libraries in a middle-sized town in the North West of England. This was complemented by a snowball strategy, where participating caregivers were asked to share the study flyer with their own social networks. Written informed consent was obtained from caregivers prior to data collection. Children received a book and caregivers received a travel reimbursement for their participation. This research has received ethical approval from the Ethics Committee of the Faculty of Science and Technology, Lancaster University (reference number: FST-2023-0791-SA-1).

Storybook Materials

Two storybooks with embedded low-frequency words were designed for this study to ensure the novelty of the storyline and that all target words were of comparable salience. Storybooks were formatted in both print and digital format so that each book served as the print condition for half of the participating dyads, and the digital condition for the other half. The size of the book was controlled across formats (single page size: 126 x 113 mm; open book / iPad screen size: 126 x 226 mm). For a detailed description of the storybook materials see (anonymised for review). Storybook materials are also available on OSF (https://osf.io/g9dw8/?view_only=baf843ea9dae49b5a45077460d6f4214) under Creative Commons Attribution 4.0 International.

The selection of the low-frequency target words was guided by the need to ensure high levels of ecological validity. Thus, we selected real words, likely to be unfamiliar to children in our age range. Our selection was informed by several criteria (see Lenhart et al., 2020 and Sarı et al., 2019 for a similar approach), including their frequency in the SUBLEX corpus of children's tv programs (van Heuven et al., 2014), and their age of acquisition (Kuperman et al., 2012). Each target word was repeated three times in the story and accompanied by an adjective on one occasion. Psycholinguistic properties of the target words and accompanying adjectives are reported in Table S7 in the Supplementary Materials.

Measures

Behavioural Coding Scheme

A coding scheme, with non-mutually exclusive codes, was developed to quantify the presence and degree of caregivers' use of verbal and gestural scaffolds in relation to target words (adapted from Hadley & Dickinson, 2019). For each target word-child combination (hereafter observation), several behaviours were coded: (1) the number of word repetitions by the caregiver in extra-textual talk; (2) whether definitional information, including synonyms, perceptual, or conceptual information, was provided by the caregiver in extra-textual talk; (3) the number of comments and questions related to the target word (e.g., "Look at that!" "Can you find the [X]?"); and (4) gestural behaviour for each observation, specifically, the presence of pointing and iconic gestures (i.e., a gesture that illustrates word meaning such as opening and closing one's hand with fingers straight to mimic a clamp). A detailed codebook is available in Table S8 Supplementary Materials. Two student assistants, blinded to our hypotheses, independently coded the video recording of the shared reading interactions after being trained with a pilot dataset. To assess inter-rater reliability, 20% of videos were double-coded. Cohen's kappa was computed for each measure. Levels of agreement ranged from *substantial* to *almost perfect*: repetition (Cohen's k = .74), definition (Cohen's k = .84), question (Cohen's k = .69), pointing (Cohen's k = .80), iconic gesture (Cohen's k = .78). The code comment was excluded from further analyses as the agreement was only moderate (Cohen's k = .51).

Caregivers' Ratings of Children's Individual Word Knowledge

A vocabulary checklist was designed after The MacArthur Communicative Development Inventories (CDI; Fenson et al., 1993). Before the shared reading activities, caregivers completed the vocabulary checklist as an indicator of their child's knowledge of target words (see Shi et al., 2022 for a similar approach). For each target word, caregivers stated whether their child understands (receptive knowledge), understands and says (receptive and expressive knowledge) the target word, or not. One point was assigned if receptive or receptive and expressive knowledge were marked as present by caregivers; otherwise, 0 was assigned.

Word learning measures

After the shared reading interactions, word learning was assessed with three tasks tapping phonological and semantic representations. A five-minute delay between the shared reading and the assessment of word learning was introduced to assess retention (in line with Hartley et al., 2020). For a detailed description of the word learning assessment protocol see (anonymised for review). A picture-naming task was developed to assess phonological recall (Adlof et al., 2021; Blewitt et al., 2009; Hargrave & Sénéchal, 2000). Each correct response for a target word was assigned 1 point. A definition task was designed to assess semantic recall, using the child-friendly procedure adopted in previous research (Blewitt & Langan, 2016). For each word, children receive 1 point for each unit of relevant information provided (e.g., object function, physical properties). For this task, inter-rater reliability was computed via an intra-class correlation (ICC) analysis with absolute agreement (ICC = .95), representing excellent agreement (Cicchetti, 1994). Finally, children were asked to identify referents of the target words in a multiple-choice task to assess semantic recognition (see Blewitt et al., 2009 and O'Toole & Kannass, 2018, for a similar approach). For each target, children received 1 point for a correct response. The chance level in this task was 25%.

Concurrent Vocabulary Knowledge

Children completed the Word Classes subtest of the Clinical Evaluation of Language Fundamentals – Preschool-2 (CELF-P2; Wiig et al., 2004). In this standardised assessment, children are shown three to four pictures (e.g., apple, shoe, bread), while the experimenter names each picture aloud. Children are instructed to identify the two words that go together best (e.g., apple and bread) and explain how they go together (e.g., both are food). Children get 1 point for each correct response, summed to provide a total vocabulary score which captures the breadth and depth of vocabulary knowledge.

Procedure

Data collection took place in an observation room that enabled non-intrusive audio and video recording of caregiver-child shared reading interactions. First, caregivers completed the vocabulary checklist. Then, caregiver-child dyads were instructed to read the two storybooks: "*I would like you to read together as you would do at home. Please take your time, I will be back when you are finished*". Each dyad

read a book presented on paper and a different book presented on an iPad, with the order of presentation counterbalanced across participants. After the shared reading activity, the child was administered the target vocabulary learning measures and the standardised vocabulary assessment.

Analytical Plan

Statistical Models

We fit Generalized Linear Mixed Models (GLMM; Baayen et al., 2008) specifying binomial error structure and logit link function for binary outcomes, and Poisson family, for count data. We included the full random effect structure supported by the data (Matuschek et al., 2017). Convergence issues were addressed by increasing the number of iterations, using different optimisers, and simplifying the random effect structure.

To investigate whether caregivers fine-tuned their communication to their children's individual lexical knowledge and explore whether book format influenced caregivers' scaffolds, we contrasted the following models with a likelihood ratio test:

- Scaffold ~ *vocabulary checklist +* prior vocabulary knowledge + age of acquisition + book format + random effect structure
- Scaffold ~ prior vocabulary knowledge + age of acquisition + book format + random effect structure

To examine whether caregivers fine-tuning supported children's word learning we contrasted the following models.

- Word learning ~ *scaffold * vocabulary checklist* + prior vocabulary knowledge + random effect structure
- Word learning ~ prior vocabulary knowledge + random effect structure

Inference Criteria

We compared our full models with null models lacking our test predictors but being otherwise identical using the likelihood ratio test to evaluate the significance of the contribution of our test predictors, whilst avoiding multiple testing (Forstmeier & Schielzeth, 2011). Beta coefficients with $p < .05$ were interpreted as significant. Null hypothesis significance testing was complemented by an examination of odds ratios (for binary outcomes), incidence rate ratios (for count outcomes) and relative confidence intervals. Marginal effects were plotted to guide the interpretation of more

complex models including interactions (Lüdecke, 2018). Significant interaction terms accompanied by improvement in the model fit, as indicated by a significant likelihood ratio test, indicated the presence of conditional effects. Marginal R^2 illustrated the proportion of variance explained by the fixed effects, while conditional R^2 illustrated the variance explained by both the fixed and random effects.

Transformations

We created a composite score for verbal and gestural scaffolds respectively. This resulted in a count variable for verbal scaffolds, where repetitions, definitions and questions were summed for each target word-child combination. For the gestural scaffold composite score, we created a binary variable which reflected the presence/absence of pointing or iconic gestures per each target word-child combination. These composites were used in our analyses to reduce issues associated with multiple testing. We provide descriptive statistics broken down by scaffold type below.

Implementation

The models were implemented in R version 4.1.3 (2022-03-10) with the function glmer of the R package lme4 (version 1.1- 33) (Bates et al., 2015). Predicted probabilities were computed using the function ggpredict of the R package ggeffects, version 1.3.2 (Lüdecke, 2018).

Results

Descriptive statistics

Across 550 observations analysed, caregivers provided no scaffold in only 10.4 % of observations, at least one scaffold in 89.6% of observations, and more than one scaffold in 75.3% of observations. The means and standard deviations for verbal and gestural scaffolds, grouped by children's lexical knowledge, are reported in Table 7. Verbal scaffolds were more frequent for unknown words than known words, and a similar pattern was observed across different types of verbal scaffolds, namely repetition, definition, and questions. There was comparable use of gestural scaffolds for unknown and known words. The means and standard deviation of verbal and gestural scaffolds grouped by book format are reported in Table 8. Means, standard deviations and correlations between quantitative variables can be consulted in Table S9 in the Supplementary Materials.

Table 7 *Means and Standard Deviations of Caregivers' Verbal and Gestural Scaffolds by Children's Individual Lexical Knowledge as Reported by Caregivers in the Vocabulary Checklist*

Scaffold	Children's individual lexical knowledge			
	Unknown		Known	
	(M)	*(SD)*	*(M)*	*(SD)*
Verbal (total)[a]	2.97	2.91	2.50	2.36
Repetitions[a]	1.32	1.51	1.11	1.18
Definitions[b]	0.51	0.50	0.33	0.47
Questions[a]	1.15	1.63	1.06	1.30
Gestural (total)[b]	0.72	0.45	0.73	0.44
Pointing[b]	0.70	0.46	0.72	0.45
Other gesture[b]	0.11	0.32	0.11	0.31

[a] count variable

[b] binary variable

Table 8 *Means and Standard Deviations of Caregivers' Verbal and Gestural Scaffolds by Book Format*

Scaffold	Book format			
	Print		Digital	
	(M)	*(SD)*	*(M)*	*(SD)*
Verbal (total)[a]	2.86	2.45	2.59	2.81
Repetitions[a]	1.30	1.33	1.12	1.37
Definitions[b]	0.43	0.50	0.40	0.49
Questions[a]	1.13	1.32	1.07	1.59
Gestural (total)[b]	0.81	0.39	0.65	0.48
Pointing[b]	0.78	0.41	0.65	0.48
Other gesture[b]	0.15	0.36	0.07	0.26

[a] count variable

[b] binary variable

Preliminary analyses were conducted to check for any possible order effects on caregivers' scaffold, both in terms of book format (print or digital presented first) and story presentation (story A or B presented first). There were no order effects on verbal scaffolds (format: β = - 0.02, *SE* = .15, *p* = .88; story: β = - 0 .25, *SE* = .16, *p* = .11), nor gestural scaffolds (format: β = .47, *SE* = .41, *p* = .25; story: β = - 0.47, *SE* = .41, *p* = .25). Preliminary analyses also identified no differences in caregivers' scaffolding as a function of child gender: verbal (β = .11, *SE* = .15, *p* = .46); gestural scaffolds (β = - 0.36, *SE* = .41, *p* = .38).

Do caregivers fine-tune their communication to their child's individual lexical knowledge?

Verbal Scaffolds

The results of the likelihood ratio test indicated that our full model was a significantly better fit to the data compared to the null model (χ^2 = 8.59, *df* = 1, *p* =

.003), supporting the significant contribution of children's individual lexical knowledge to caregivers' verbal scaffolds. Our full model suggests that caregivers calibrated their verbal scaffolds and provided significantly fewer scaffolds for words that they thought their children already knew (β = -0.19, *CI* = - 0.32 – -0.07, *p* = 0.003). Importantly, the model controlled for children's broader vocabulary knowledge, which was not significant (β = 0.01, *CI* = -0.05 – 0.07, *p* = 0.672), and for target words' age of acquisition, which was no longer significant once children's individual lexical knowledge was included in the model (β = 0.04, *CI* = -0.01 – 0.08, *p* = 0.099). Results also revealed that caregivers provided significantly fewer verbal scaffolds when reading a digital compared to a print book (β = - 0.13, *CI* = -0.23 – -0.02, *p* = 0.017). Model results are reported in Table 9. Incidence rate ratios can be consulted in Table S10 in the Supplementary Materials.

Table 9 *Effects of Individual Lexical Knowledge (Vocabulary Checklist) on Caregivers' Verbal Scaffolds*

	Verbal scaffolds							
	Null model				Full model			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	0.12	0.36	-0.57 – 0.82	0.726	0.53	0.38	-0.22 – 1.28	0.169
Vocabulary knowledge	0.01	0.03	-0.05 – 0.07	0.700	0.01	0.03	-0.05 – 0.07	0.672
Age of acquisition	0.07	0.02	0.03 – 0.10	**<0.001**	0.04	0.02	-0.01 – 0.08	0.099
Book format [digital]	-0.14	0.05	-0.24 – -0.03	**0.009**	-0.13	0.05	-0.23 – -0.02	**0.017**
Individual lexical knowledge					-0.19	0.07	-0.32 – -0.07	**0.003**
Random Effects								
σ^2	0.36				0.36			
τ_{00}	0.42 $_{child\_ID}$				0.43 $_{child\_ID}$			
ICC	0.54				0.54			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.021 / 0.548				0.029 / 0.555			

Gestural Scaffolds

The likelihood ratio test did not support the significant contribution of children's individual lexical knowledge to caregivers' gestural scaffolds (χ^2 = 0.846, *df* = 1, *p* = 0.358). In contrast to the results reported above for verbal scaffold, there was no evidence that caregivers modulated their gestural scaffolds for target words according to their perception of their children's individual word knowledge (β = 0.27, *CI* = -0.31 – 0.85, *p* = 0.359). As in our previous analysis, the control predictors of children's general vocabulary knowledge and target words' age of acquisition were not significant (β = -0.01, *CI* = -0.17 – 0.15, *p* = 0.870; β = 0.04, *CI* = -0.14 – 0.23, *p* = 0.643, respectively). In line with the results for verbal scaffolds, there was evidence that caregivers provided significantly fewer gestural scaffolds when reading a digital compared to a print book (β = -1.20, *CI* = -1.68 – -0.71, *p* < 0.001). Model results are reported in Table 10. Odds ratios can be consulted in Table S11 in the Supplementary Materials.

Table 10 *Effects of Individual Lexical Knowledge (Vocabulary Checklist) on Caregivers' Gestural Scaffolds*

	Gestural scaffolds							
	Null model				Full model			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	2.24	1.15	-0.01 – 4.49	0.051	1.67	1.30	-0.88 – 4.23	0.200
Vocabulary knowledge	-0.01	0.08	-0.17 – 0.15	0.882	-0.01	0.08	-0.17 – 0.15	0.870
Age of acquisition	-0.00	0.08	-0.16 – 0.15	0.955	0.04	0.09	-0.14 – 0.23	0.643
Book format [digital]	-1.17	0.25	-1.65 – -0.69	**<0.001**	-1.20	0.25	-1.68 – -0.71	**<0.001**
Individual lexical knowledge					0.27	0.30	-0.31 – 0.85	0.359
Random Effects								
σ^2	3.29				3.29			
τ_{00}	2.52 $_{child\_ID}$				2.54 $_{child\_ID}$			
ICC	0.43				0.44			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.056 / 0.465				0.058 / 0.468			

Does Fine-Tuning Support Children's Word Learning?

As stated in our pre-registration, the effect of fine-tuning on word learning was examined only for scaffolds where there was evidence of fine-tuning. Our previous analyses indicated that caregivers fine-tune their verbal but not gestural scaffolds. Therefore, we analysed the contribution of *verbal* fine-tuning to word learning. The effect of book format on word learning was examined as a separate research question in a partially overlapping dataset (anonymised for review). Results of these analyses revealed no significant main effect of book format on word learning. For this reason, book format is not included in the following analyses. Word learning was evidenced by children's performance in three tasks: phonological recall, semantic recall, and semantic recognition. The results for each outcome are reported below.

Phonological Recall

The likelihood ratio test indicated that our full model was a better fit to the data than our null model (χ^2 = 166.72, *df* = 3, *p* < 0.001), but the coefficients of the full model (Table 4) indicated that the interaction between individual lexical knowledge and verbal scaffolds was not significant (β = 0.09, *CI* = -0.23 – 0.40, *p* = 0.597), suggesting no clear benefit of fine-tuning for phonological recall. Vocabulary knowledge was significantly and positively related to phonological recall (β = 0.11, *CI* = 0.02 – 0.19, *p* = 0.014), and the same was true for individual lexical knowledge (β = 3.07, *CI* = 2.06 – 4.08, *p* < 0.001). Model results are reported in Table 11. Odds ratios can be consulted in Table S12 in the Supplementary Materials.

Table 11 *Effects of Fine-Tuning on Phonological Recall*

	Phonological recall							
	Null model				Full model			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-1.85	0.41	-2.66 – -1.04	**<0.001**	-4.14	0.68	-5.47 – -2.80	**<0.001**
Vocabulary knowledge	0.08	0.04	0.01 – 0.16	**0.031**	0.11	0.04	0.02 – 0.19	**0.014**
Individual lexical knowledge					3.07	0.52	2.06 – 4.08	**<0.001**
Verbal scaffolds					-0.13	0.15	-0.44 – 0.17	0.382
Individual lexical knowledge × Verbal scaffolds					0.09	0.16	-0.23 – 0.40	0.597
Random Effects								
σ^2	3.29				3.29			
τ_{00}	0.00 $_{child\_ID}$				0.00 $_{child\_ID}$			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.014 / NA				0.473 / NA			

Semantic Recall

The likelihood ratio test indicated that our full model was a better fit to the data than our null model (χ^2 = 160.46, *df* = 3, *p* < 0.001), and the coefficients of the full model (Table 5) supported the presence of a significant interaction between individual lexical knowledge and verbal scaffolds (β = -0.08, *CI* = -0.15 – -0.01, *p* = 0.024). Verbal scaffolds predicted semantic recall only for words that caregivers thought their children did not know, supporting the contribution of caregivers' fine-tuning to children's semantic recall. Paralleling our previous analysis on phonological recall, vocabulary knowledge was significantly and positively related to semantic recall (β = 0.12, *CI* = 0.05 – 0.19, *p* = 0.001), and the same applied to individual lexical knowledge (β = 1.61, *CI* = 1.28 – 1.95, *p* < 0.001). Model results are reported in Table 12. Incidence rate ratios can be consulted in Table S13 in the Supplementary Materials.

Table 12 *Effects of Fine-Tuning on Semantic Recall*

	Semantic recall							
	Null model				Full model			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-1.72	0.42	-2.53 – -0.90	**<0.001**	-2.76	0.42	-3.59 – -1.93	**<0.001**
Vocabulary knowledge	0.13	0.04	0.05 – 0.20	**0.001**	0.12	0.04	0.05 – 0.19	**0.001**
Individual lexical knowledge					1.61	0.17	1.28 – 1.95	**<0.001**
Verbal scaffolds					0.07	0.03	0.01 – 0.13	**0.033**
Individual lexical knowledge × Verbal scaffolds					-0.08	0.04	-0.15 – -0.01	**0.024**
Random Effects								
σ^2	0.93				0.93			
τ_{00}	0.48 $_{child\_ID}$				0.43 $_{child\_ID}$			
ICC	0.34				0.32			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.074 / 0.392				0.309 / 0.528			

Semantic Recognition

The likelihood ratio test indicated that our full model was a better fit to the data than our null model (χ^2 = 32.18, *df* = 3, *p* < 0.001), yet the coefficients of the full model (Table 6) did not support the presence of a significant interaction between individual lexical knowledge and verbal scaffolds (β = 0.07, *CI* = -0.12 – 0.26, *p* = 0.484). In line with previous analyses, vocabulary knowledge was significantly and positively related to semantic recognition (β = 0.18, *CI* = 0.07 – 0.29, *p* = 0.002), and the same applied to individual lexical knowledge (β = 1.10, *CI* = 0.45 – 1.76, *p* < 0.001). Model results are reported in Table 13. Odds ratios can be consulted in Table S14 in the Supplementary Materials.

Table 13 *Effects of Fine-Tuning on Semantic Recognition*

	Semantic recognition							
	Null model				Full model			
Predictors	*Estimate*	*SE*	*CI*	*p*	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-0.53	0.54	-1.59 – 0.52	0.321	-1.23	0.61	-2.42 – -0.04	**0.043**
Vocabulary knowledge	0.17	0.05	0.07 – 0.27	**0.001**	0.18	0.06	0.07 – 0.29	**0.002**
Individual lexical knowledge					1.10	0.33	0.45 – 1.76	**0.001**
Verbal scaffolds					0.02	0.06	-0.09 – 0.13	0.716
Individual lexical knowledge × Verbal scaffolds					0.07	0.10	-0.12 – 0.26	0.484
Random Effects								
σ^2	3.29				3.29			
τ_{00}	0.58 $_{child\_ID}$				0.72 $_{child\_ID}$			
ICC	0.15				0.18			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.051 / 0.194				0.143 / 0.298			

Discussion

The present study investigated whether and how caregivers display fine-tuning behaviours during shared reading interactions and whether this supports children's word learning. We also explored whether the book presentation format affected caregivers' scaffolding. Caregivers adjusted their communication as a function of their children's individual lexical knowledge. This was apparent for verbal but not gestural scaffolds. There was evidence that fine-tuning supported word learning as assessed with a semantic recall task. Finally, caregivers provided significantly fewer verbal and gestural scaffolds when sharing a digital compared to a print book. We discuss each of these findings and their theoretical and practical implications in the following sections.

Replicating and extending prior research (Donnellan et al., 2023; Leung et al., 2021; Shi et al., 2022), these findings provide the first empirical evidence that caregivers fine-tune their verbal behaviour during shared book reading interactions. By demonstrating that fine-tuning behaviour extends beyond tablet-based reference game and toy play interactions (Donnellan et al., 2023; Leung et al., 2021; Shi et al., 2022), this study supports the generalisability of fine-tuning hypothesis. This has important theoretical implications as it provides further clarity on caregivers' ability to precisely tailor their language to their children's lexical knowledge. Given the central role of shared book reading for early vocabulary development, these findings have also important practical implications. Adapting support to children's individual word knowledge during shared book reading may be a promising way to reduce their cognitive load and optimise learning across a range of ability levels (Sadoski & Lawrence, 2023).

These results also demonstrate for the first time that caregivers display fine-tuning behaviour with 4- to 5-year-olds, an older age group compared to those examined in prior research (Donnellan et al., 2023; Leung et al., 2021; Shi et al., 2022). Whilst certain properties of child-directed language have been documented to change as children age (Huttenlocher et al., 2010), our data suggest that caregivers fine-tuning is a robust mechanism, adopted by caregivers for an extended period in child development, spanning from toddlerhood, as demonstrated in previous work (Leung et al., 2021), until, at least, five years of age, as illustrated in the current study.

In contrast to Donnellan and colleagues (2023) we did not find evidence that caregivers modulated their gestures as a function of children's lexical knowledge. There are different possible interpretations for these contradictory findings. First, it should be noted that pointing and iconic gestures were treated as the same gestural composite in our analyses to minimise the false discovery rate associated with multiple testing. Donnellan and colleagues (2023) differentiated between pointing and representational gestures and found evidence of caregivers' modulation for pointing but not for representational gestures. It is possible that fine-tuning is operational only for certain kinds of gestures, which would make our composite measure unsuitable for studying fine-tuning. Yet our descriptive statistics do not support such an interpretation, as they suggest similar levels of both pointing and iconic gestures across known and unknown words. Thus, an alternative possibility is that pointing was particularly important in the toy-play interactions as three unknown word-object associations were presented simultaneously (Donnellan et al., 2023). In contrast, in our shared reading materials unknown word-object associations were presented one page at a time, making the identification of the referent less challenging. Future studies investigating different learning and interactional contexts with varying degrees of referential ambiguity are needed to clarify the conditions under which caregivers modulate their use of gestures.

We followed up on the relation between the identified verbal fine-tuning and children’s word learning. Our study was the first to include three different measures of word learning to capture complementary aspects of phonological and semantic representations. Results suggest that fine-tuning predicted semantic recall, but not phonological recall or semantic recognition. Whilst the relation between fine-tuning and word learning, in general, is in line with prior research (Donnellan et al., 2023; Leung et al., 2021; Shi et al., 2022), the specific measures of word learning for which this relation was found differ: semantic recall in the current study vs semantic recognition in prior research. This difference could be explained, at least in part, by the inclusion of older children in our study. For our age group and learning context, the benefit of fine-tuning may have become apparent only when assessing children’s deeper understanding of word meaning through semantic recall. Indeed, prior work also suggests that the positive effect of a scaffolding-like procedure during shared

reading becomes evident only when considering semantic recall (Blewitt et al., 2009). From a theoretical standpoint, the identification of a specific link between verbal fine-tuning and semantic recall is in line with the notions of scaffolding (Wood et al., 1976) and zone of proximal development (Vygotsky, 1978). By taking into account children's individual lexical knowledge, caregivers provide just the right level of support that enables children to efficiently acquire the semantic features of new words encountered during shared reading.

Caregivers were less likely to provide both verbal and gestural scaffolds when reading a digital compared to a print book, which replicates and extends prior research on interaction quality during print-based and digital-based shared reading (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020). These results also align with recent work underscoring the importance of considering the conversational context (Brinchmann et al., 2023) and features of books (O'Rear et al., 2023) when examining the properties of caregivers' input. Considering that early literacy experiences are rapidly changing, and digital books are becoming increasingly common (Kucirkova, 2019), these findings are unique in contributing to the evidence-base needed to inform recommendations for users of research.

One means through which scaffolds might enhance learning is through a reduction in overall cognitive load. Scaffolds may reduce cognitive load through different mechanisms. One possible mechanism is directing and focusing learners' attention on relevant features in the environment. Pointing and repetition, for example, help to direct and focus attention to an object or picture and its name. Another possible mechanism through which scaffolds may reduce cognitive load is through the provision of multimodal cues (Sadoski & Lawrence, 2023). Gestures, shown to be effective in older children's novel verb learning (De Nooijer et al., 2013), provide additional multimodal support, which is particularly important from an embodied cognition perspective (Smith & Gasser, 2005). Questions and definition-based expansions would reduce the inferential load on the child, an important predictor of older children's word learning from text (Cain, Lemmon, et al., 2004). Future research might usefully seek to disentangle the contributions of scaffolds that direct attention from those that provide additional multimodal support for

phonological and semantic representations of the word, to understand better the mechanisms that underpin different aspects of word learning.

Limitations and Future Directions

It should be acknowledged that this study comes with limitations and suggestions for further research, additionally to those already noted. It is important to note that a relatively small set of words was included in our study materials. We discussed why this is a common issue in the shared reading literature and suggested potential solutions in (anonymised). Relatedly, only concrete nouns for animals and tools were studied in this work. It is critical for future research to consider a range of word types, including adjectives and verbs, to establish if these results generalise for words other than concrete nouns. The present study provides the first evidence that fine-tuning generalises to the shared reading context, but future work is needed to directly compare fine-tuning in different interactional contexts (e.g., toy play, shared reading, mealtime conversations) to identify similarities and differences in use and effectiveness across contexts. Also, we did not analyse prosody, another type of scaffold that caregivers adapt in relation to 3- to 4-year-olds' individual lexical knowledge, and shown to predict word learning during toy play interactions, and vocabulary one year later (Shi et al., 2022). Such work would provide much-needed information to inform our understanding of how fine-tuning operates at the individual level, across contexts and scaffolds.

It should also be noted that current paradigms to investigate fine-tuning rely on caregivers' reports of children's word knowledge. There are at least two limitations associated with this approach. One is the reliability of these measures, and another one is that asking caregivers to reflect on children's lexical knowledge may influence their subsequent behaviour. Advances in automated analyses of the home environment may provide a useful tool to overcome this limitation. This study is also limited in that it measures word learning over a short timescale. Future research should consider the effect of fine-tuning over longer time scales, for instance, by including longitudinal measures of vocabulary development (Shi et al., 2022). Importantly, these measures should capture both the breadth and depth of vocabulary knowledge to reflect the richness of the developing lexicon (Hadley et al., 2019).

Our findings are also limited to the cultural context of our participating sample, which was recruited within travelling distance of our child assessment labs. Therefore, replication to different cultural contexts is warranted. Finally, our identification of the detrimental effects of digital format on caregivers' multimodal scaffolds underscores the importance of conducting further research into the design of digital books. Understanding which features of digital books can enhance caregiver-child interactions and support children's understanding and learning is a pressing challenge in an increasingly digital world.

Conclusions

The importance of caregivers' mediation to enhance word learning from shared reading is well established. Here we showed for the first time that caregivers adapt their verbal communication to their children's individual lexical knowledge during shared reading and that this supports children's learning of new word meanings. These results suggest that caregivers' mediation is particularly effective when it is tailored and individualised to the child's current knowledge of specific words. We also identified the detrimental effects of digital book format on caregivers' multimodal scaffolds, urging further research on how best to design digital books to support social interaction, understanding and learning in early childhood. This area of inquiry may be fruitfully informed by caregivers' fine-tuning behaviours.

Acknowledgements

We sincerely thank all the participating caregivers and children. We are grateful to Simya Aravamuthan, Gracey Caller, and Arwen Hon for their support with data scoring, inputting, and recruitment of participants. We thank Phoebe Schaw and Ffion Jones for their support with coding video recordings of shared reading interactions. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Actions grant agreement No 857897.

Open Research

Anonymised data and code necessary to reproduce the analyses presented here are publicly accessible, as are the materials necessary to attempt to replicate the findings. Our research questions, hypotheses, design, and analytic plan were pre-

registered after data collection and descriptive statistics were performed on a subsample (*n* = 42).

Supporting Information

Data, code, materials, and the preregistration for this research are available on the project's OSF repository (OSF; https://osf.io/g9dw8/?view_only=baf843ea9dae49b5a45077460d6f4214 https://osf.io/4cm73?view_only=37ff877cb13a491bb424e2b07aebb0a5).

CRediT Author Statement

Laura Diprossimo**:** Conceptualization, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. Kate Cain**:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

References

Adlof, S. M., Baron, L. S., Bell, B. A., & Scoggins, J. (2021). Spoken Word Learning in Children With Developmental Language Disorder or Dyslexia. *Journal of Speech, Language, and Hearing Research*, *64*, 2734–2749. https://doi.org/10.23641/asha

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barnes, E. M., Hadley, E. B., Lawson-Adams, J., & Dickinson, D. K. (2023). Nonverbal supports for word learning: Prekindergarten teachers' gesturing practices during shared book reading. *Early Childhood Research Quarterly*, *64*, 302–312. https://doi.org/10.1016/j.ecresq.2023.04.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1 SE-Articles), 1–48. https://doi.org/10.18637/jss.v067.i01

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology*, *150*, 404–410. https://doi.org/10.1016/j.jecp.2016.06.009

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared Book Reading: When and How Questions Affect Young Children's Word Learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Brinchmann, E. I., Røe-Indregård, H., Karlsen, J., Schauber, S. K., & Hagtvet, B. E. (2023). The linguistic complexity of adult and child contextualized and decontextualized talk. *Child Development*, *94*(5), 1368–1380. https://doi.org/10.1111/cdev.13932

Cain, K., Lemmon, K., & Oakhill, J. (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. https://doi.org/10.1037/0022-0663.96.4.671

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *4*, 284–290.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica*, *144*(1), 173–179. https://doi.org/10.1016/j.actpsy.2013.05.013

Donnellan, E., Jordan-Barros, A., Theofilogiannakou, N., Brekelmans, G., Murgiano, M., Motamedi, Y., Grzyb, B., Gu, Y., & Vigliocco, G. (2023). The impact of caregivers' multimodal behaviours on children's word learning: A corpus-based investigation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 693–701.

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). The MacArthur Communicative Development Inventories: User's guide and technical manual. In *San Diego, CA: Singular Publishing Group.*

Fernald, A., & Mazzie, C. (1991). Prosody and Focus in Speech to Infants and Adults. *Developmental Psychology*, *27*(2), 209–221. https://doi.org/10.1037/0012-1649.27.2.209

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The Effects of Shared Storybook Reading on Word Learning: A Meta-Analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/http://dx.doi.org/10.1037/dev0000512

Flack, Z. M., & Horst, J. S. (2018). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development*, *27*(1), 1–12. https://doi.org/10.1002/icd.2047

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Fusaroli, R., Weed, E., Rocca, R., Fein, D., & Naigles, L. (2023). Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development. *Cognition*, *236*. https://doi.org/10.1016/j.cognition.2023.105422

Government of the UK. (2019). *English indices of deprivation 2019*. https://imd-by-postcode.opendatacommunities.org/imd/2019

Government of the UK. (2022, December). *Regional ethnic diversity*. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest/

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, *18*(5), 797–824. https://doi.org/10.1111/infa.12006

Hadley, E. B., & Dickinson, D. K. (2019). Cues for word-learning during shared book-reading and guided play in preschool. *Journal of Child Language*, *46*(6), 1202–1227. https://doi.org/10.1017/S0305000919000552

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Building Semantic Networks: The Impact of a Vocabulary Intervention on Preschoolers' Depth of Word Knowledge. *Reading Research Quarterly*, *54*(1), 41–61. https://doi.org/10.1002/rrq.225

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*, 104265. https://doi.org/10.1016/j.cognition.2020.104265

Hills, T. (2013). The company that words keep: Comparing the statistical structure of child- Versus adult-Directed language. *Journal of Child Language*, *40*(3), 586–604. https://doi.org/10.1017/S0305000912000165

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Korat, O., & Or, T. (2010). How new technology influences parent-child interaction: The case of e-book reading. *First Language*, *30*(2), 139–154. https://doi.org/10.1177/0142723709359242

Kucirkova, N. (2019). Children's Reading With Digital Books: Past Moving Quickly to the Future. *Child Development Perspectives*, *13*(4), 208–214. https://doi.org/10.1111/cdep.12339

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*(4), 978–990.

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2019). The effects of questions during shared-reading: Do demand-level and placement really matter? *Early Childhood Research Quarterly*, *47*, 49–61. https://doi.org/10.1016/j.ecresq.2018.10.006

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2020). More than words: Narrator engagement during storytelling increases children's word learning, story comprehension, and on-task behavior. *Early Childhood Research Quarterly*, *51*, 338–351. https://doi.org/10.1016/j.ecresq.2019.12.009

Leung, A., Tunkel, A., & Yurovsky, D. (2021). Parents Fine-Tune Their Speech to Children's Vocabulary Knowledge. *Psychological Science*, *32*(7), 975–984. https://doi.org/10.1177/0956797621993104

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant-and adult-directed speech. *Language Learning and Development*, *7*(3), 185–201. https://doi.org/10.1080/15475441.2011.579839

Masek, L. R., McMillan, B. T. M., Paterson, S. J., Tamis-LeMonda, C. S., Golinkoff, R. M., & Hirsh-Pasek, K. (2021). Where language meets attention: How contingent interactions promote learning. *Developmental Review*, *60*. https://doi.org/10.1016/j.dr.2021.100961

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Ministry of Housing Communities and Local Government. (2019). *The English Indices of Deprivation 2019: Research Report*. https://assets.publishing.service.gov.uk/media/5d8b364ced915d03709e3cf2/IoD2019_Research_Report.pdf

Mol, S. E., & Bus, A. G. (2011). To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. (2019). Parent-Toddler Social Reciprocity during Reading from Electronic Tablets vs Print Books. *JAMA Pediatrics*, *173*(11), 1076–1083. https://doi.org/10.1001/jamapediatrics.2019.3480

Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*(1), 43–79. https://doi.org/10.1017/S0305000900005584

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*(3), 423–430. https://doi.org/10.1016/j.cognition.2008.10.004

O'Rear, C. D., Seip, I., Azar, J., Baroody, A. J., & McNeil, N. M. (2023). Features in children's counting books that lead dyads to both count and label sets during shared book reading. *Child Development*, *94*(4), 985–1001. https://doi.org/10.1111/cdev.13915

O'Toole, K. J., & Kannass, K. N. (2018). Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology*, *173*, 100–115. https://doi.org/10.1016/j.jecp.2018.03.013

Ozturk, G., & Hill, S. (2020). Mother–child interactions during shared reading with digital and print books. *Early Child Development and Care*, *190*(9), 1425–1440. https://doi.org/10.1080/03004430.2018.1538977

Richter, A., & Courage, M. L. (2017). Comparing electronic and paper storybooks for preschoolers: Attention, engagement, and recall. *Journal of Applied Developmental Psychology*, *48*, 92–102. https://doi.org/10.1016/j.appdev.2017.01.002

Sadoski, M., & Lawrence, B. (2023). Abstract Vocabulary Development: Embodied Theory and Practice. In *Educational Psychology Review* (Vol. 35, Issue 3). Springer. https://doi.org/10.1007/s10648-023-09802-9

Sarı, B., Başal, H. A., Takacs, Z. K., & Bus, A. G. (2019). A randomized controlled trial to test efficacy of digital enhancements of storybooks in support of narrative comprehension and word learning. *Journal of Experimental Child Psychology*, *179*, 212–226. https://doi.org/10.1016/j.jecp.2018.11.006

Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children's word learning. *Developmental Psychology*, *52*(6), 879–886. https://doi.org/10.1037/dev0000125

Sénéchal, M., Pagan, S., Lever, R., & Ouellette, G. P. (2008). Relations among the frequency of shared reading and 4-year-old children's vocabulary, morphological and syntax comprehension, and narrative skills. *Early Education and Development*, *19*(1), 27–44. https://doi.org/10.1080/10409280701838710

Shi, J., Gu, Y., & Vigliocco, G. (2022). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*. https://doi.org/10.1111/desc.13357

Smith, L., & Gasser, M. (2005). The Development of Embodied Cognition: Six Lessons from Babies. *Artificial Life*, *11*(1–2), 13–29.

Snow, C. E. (1972). Mothers' Speech to Children Learning Language. *Child Development*, *43*(2), 549–565.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. In *Harvard University Press*.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Wainwright, B. R., Allen, M. L., & Cain, K. (2020). Narrative comprehension and engagement with e-books vs. paper-books in autism spectrum condition. *Autism and Developmental Language Impairments*, *5*. https://doi.org/10.1177/2396941520917943

Wiig, E. H. , Secord, W. A. , & Semel, E. (2004). Clinical evaluation of language fundamentals—Preschool, (CELF Preschool-2). In *Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.*

Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Chapter 4: The Associations between Child and Item Characteristics, Use of Vocabulary Scaffold, and Reading Comprehension in a Digital Environment: Insights from a Big Data Approach

Linking Statement

This pre-registered secondary data analysis (Project https://osf.io/hfe53/; Registration https://doi.org/10.17605/OSF.IO/62C4Q) leveraged on an existing big dataset of user log files, resulting from children's interaction with a digital reading supplement. This study was necessitated by the travel restrictions during 2020-2021 and starting the PhD remotely. It focuses on a slightly older age group than the participants in chapters 2 and 3, namely, children aged 5 to 8 years who were learning to read. It also focuses entirely on digital reading, by exploring the links between the use of vocabulary scaffolds and reading comprehension in a digital reading supplement. This study addresses themes common to the earlier work, by looking at the role of scaffolds and the influence of child characteristics on their uptake and effect.

Author Contribution Statement

In the chapter entitled "The associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment: Insights from a big data approach" the authors agree to the following contributions:

Laura Diprossimo, 60% (Conceptualization, Formal analysis, Data curation, Writing – original draft, Visualization, Project administration)

Anastasia Ushakova, 12.5% (Conceptualization, Formal analysis, Writing – original draft, Project administration)

Jennifer Zoski, 5% (Resources, Writing – review & editing)

Harrison Gamble, 5% (Resources, Data curation, Writing – review & editing)

Robin Irey, 5% (Writing – review & editing)

Kate Cain, 12.5% (Conceptualization, Writing – original draft, Supervision, Funding acquisition)

Text as it appears in Diprossimo, L., Ushakova, A., Zoski, J., Gamble, H., Irey, R., & Cain, K. (2023). The associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment: Insights from a big data approach. *Contemporary Educational Psychology*, 73, 102165. https://doi.org/10.1016/j.cedpsych.2023.102165

Abstract

Scaffolding features that provide multimodal support for the pronunciation and meaning of words are increasingly common in digital reading environments. These vocabulary scaffolds are intended to aid the accurate pronunciation and understanding of individual words in context, thus supporting both vocabulary development and comprehension of text. However, the evidence on their efficacy remains inconclusive. The present study adds to the evidence base by examining: 1) whether child characteristics predict the use of vocabulary scaffolds; 2) whether the use of vocabulary scaffolds is associated with reading comprehension performance; and 3) whether the association between the use of scaffolds and reading comprehension is modulated by child and/or item characteristics. A large cohort (N ~ 120,000) of 5- to 8-year-old children in the United States interacted with a gamified digital reading environment with embedded vocabulary scaffolds, thereby generating a large observational dataset of user log files. Confirmatory analyses with Generalized Linear Mixed Models (GLMMs) indicated that children with lower literacy skills, beginning readers, girls, and bilingual students were more likely to use the scaffold. Overall, the use of scaffolds was associated with better reading comprehension performance. The association between the use of scaffolds and reading comprehension was modulated by both child and item characteristics. We conclude that vocabulary scaffolds may be promising tools to facilitate reading comprehension and reduce performance differences amongst diverse learners in digital reading environments. Educational implications and recommendations for future research are discussed.

Keywords: vocabulary, reading comprehension, scaffolding features, multimedia learning, big data

Introduction

Digital texts afford novel opportunities to scaffold children's reading comprehension (Dalton et al., 2011; Gonzalez, 2014; Proctor et al., 2007; Zou et al., 2021). They provide a critical opportunity for learning through in-built scaffolding features that offer multimodal support (i.e., auditory and visual) for the pronunciation and meaning of words. These features, also referred to in the literature as hypertext glosses (Yun, 2011), dictionaries (Furenes et al., 2021), annotations (Akbulut, 2007) and vocabulary supports (Gonzalez, 2014; Proctor et al., 2007), are intended to support the accurate pronunciation and understanding of individual words in context, potentially promoting both vocabulary development and comprehension of the text. To date, evidence supporting the benefits of digital scaffolds remains elusive, and a thorough evaluation of these features is lacking. First, little is known about which child characteristics predict the use of such scaffolds when they are available to the reader on an as-needed basis. Second, the association between use of scaffolds and reading comprehension remains unclear, with mixed evidence arising potentially from diverse participants and item characteristics across studies (Abraham, 2008; Furenes et al., 2021).

Our study was designed to address these gaps in our knowledge to provide a detailed and critical evaluation of the associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment. We exploited a large observational dataset of children's interactions with a real-world digital reading product that was supplementary to the established classroom curriculum. Our analyses identified critical factors that were associated with the use of scaffolds, and which modulated the associations between the use of scaffolds and reading comprehension.

Scaffolding Principle

A scaffold was originally defined as a process that enables a child or novice to solve a problem, carry out a task, or achieve a goal that is beyond their unassisted efforts (Wood et al., 1976). Contemporary definitions of scaffolds are, however, more nuanced and include several tools and aids. Scaffolds have been investigated across a wide range of learning contexts, including instruction in narrative skills (Pesco & Gagné, 2017), problem-based learning for science, technology, engineering and

mathematics education (N. J. Kim et al., 2018), and simulation-based learning in medical or teacher training (Chernikova et al., 2020). Scaffolds have been found to benefit learning in different age groups, including children (Pesco & Gagné, 2017) and adults (Chernikova et al., 2020).

With respect to reading comprehension, instructional scaffolding represents a flexible and adaptable model (Clark & Graves, 2005). A scaffolded reading experience may involve a range of activities that can take place before, during, or after reading. These activities include pre-teaching vocabulary, relating the content to children's lives, asking questions and discussing the text (Clark & Graves, 2005). These scaffolds have all proved beneficial for reading comprehension (Blything et al., 2020; Degener & Berne, 2017; Elleman et al., 2009; Mckeown et al., 2009; Zucker et al., 2010). In relation to the specific focus of this study, a meta-analysis evaluating the effect of vocabulary instruction on passage-level comprehension found a positive impact for classroom-based instruction focused on increasing word knowledge (Elleman et al., 2009). This provides strong evidence for the potential benefit of scaffolds for vocabulary items to support understanding of text.

Beyond traditional instructional scaffolding, technology-based scaffolding is becoming increasingly common. Digital environments, and consequently digital texts, afford novel opportunities to scaffold children's reading comprehension (Dalton et al., 2011; Gonzalez, 2014; Proctor et al., 2007; Zou et al., 2021). Indeed, the medium lends itself to support independent reading with built-in scaffolds. For instance, empirical work has shown that text-to-speech read aloud functionality scaffolds text comprehension for grade 3 and 4 students with disabilities (Gonzalez, 2014). Moreover, an intervention study found that the use of scaffolds that provide comprehension strategies embedded in digital texts positively relates to reading comprehension gains in 4th graders (Proctor et al., 2007). Importantly, digital environments often incorporate gamification and adaptivity principles, which are both powerful tools for promoting motivation and learning (Cohen et al., 1982; Lämsä et al., 2018; Manzano-León et al., 2021). Despite the promise of digital supplements in education, there is consensus that specific features need to be evaluated carefully in relation to learning outcomes to disentangle which features works best and for whom (Lämsä et al., 2018; Manzano-León et al., 2021; Mayer, 2020). Our focus here is on the

evaluation of scaffolding features which were designed to promote understanding of single words in context and embedded in a gamified digital reading environment. The scaffolds included written definitions, and auditory and visual support for pronunciation and meaning of target words, which could be used on an as-needed basis. A comprehensive evaluation of these features is crucial given the strong and reciprocal relationship between vocabulary knowledge and reading comprehension (Cain & Oakhill, 2018).

Vocabulary Knowledge and Reading Comprehension

Vocabulary knowledge is a unique predictor of word reading in kindergarteners (Y. S. Kim et al., 2014), and reading comprehension in later grades (Oakhill & Cain, 2012; Quinn et al., 2015). As noted above, a meta-analysis exploring the impact of vocabulary instruction on passage-level comprehension for school-age children found a positive impact for this type of scaffold, and indicated that the impact was three times greater for students with reading difficulties (Elleman et al., 2009). Thus, vocabulary instruction is viewed as a crucial ingredient of interventions targeting reading fluency and reading comprehension (Wolf & Katzir-Cohen, 2009). Importantly, vocabulary development and reading comprehension, and its core skills such as inference, are interdependent (Cain & Oakhill, 2018; Language and Reading Research Consortium et al., 2019): vocabulary knowledge is both foundational to reading comprehension, and also boosted by reading practice, presumably because good text comprehension skill enables readers to derive the meanings of unfamiliar words from context (Cain et al., 2003). Written text indeed provides a rich source of exposure to a range of less frequent and less familiar words in context (Cunningham & Stanovich, 1998; Montag et al., 2015).

Vocabulary Scaffolding Features

Given both the strong relationship between vocabulary and reading comprehension (Cain & Oakhill, 2018), and written text itself as a rich source of new word learning (Cunningham & Stanovich, 1998; Montag et al., 2015), there is a need to understand the role of vocabulary scaffolds in digital texts. These scaffolds have proven beneficial for vocabulary learning for struggling readers and second language learners (Liu & Leveridge, 2017; Proctor et al., 2007; Yun, 2011), as well as for young children in the context of shared reading (Furenes et al., 2021). Indeed, inferring the

meanings of novel words from linguistic contexts may be particularly challenging for children with poor reading comprehension (Cain et al., 2003) and for second language learners (Nassaji, 2003).

The positive effect of vocabulary scaffolding features that combine pictorial support and definitions on word learning aligns with the wealth of research documenting the benefits of studying vocabulary with verbal explanations and images as compared to studying with only verbal explanations (Akbulut, 2007; Andrä et al., 2020; Hald et al., 2016; D. Kim & Gilman, 2008; Rowe et al., 2013; Tonzar et al., 2009). This so-called picture superiority effect in vocabulary acquisition has been explained by the dual coding theory (Paivio & Csapo, 1973), which posits that image and verbal memory codes are independent and have an additive effect on recall. Similarly, the cognitive theory of multimedia learning explains the picture superiority effect as a multimedia effect, in which students learn more deeply from words and pictures than from words alone (Mayer, 2017; Mayer, 2003; Mayer & Moreno, 2003; Mayer & Sims, 1994).

Evidence documenting the impact of vocabulary scaffolds on reading comprehension is mixed. Whilst some meta-analytic evidence suggests that vocabulary scaffolds have an overall medium size effect on second language reading comprehension in university students (Abraham, 2008), other meta-analytic research shows that dictionaries embedded in digital storybooks aid word learning, but have a negligible or negative impact on story comprehension in children up to 8 years of age (Furenes et al., 2021). However, these results refer to a small subsample (n = 5) of the studies included in this meta-analysis. Indeed, another study of children aged 4 to 6 (Korat & Shamir, 2012), not included in the synthesis by Furenes and colleagues (2021), found a significant and positive association between story comprehension and gains in knowledge and decoding of words that received direct computer support (i.e., support was given a priori to all children for target words) but did not find gains in knowledge and decoding of words that did not receive the computer supports. This finding suggests a positive association between the provision of vocabulary scaffolds and young children's story comprehension in digital environments. Studies reporting a positive association between use of vocabulary scaffolding features and reading comprehension are in line with both dual coding theory (Paivio & Csapo, 1973) and the

cognitive theory of multimedia learning (Mayer, 2017; Mayer, 2003; Mayer & Moreno, 2003; Mayer & Sims, 1994), and also with research involving older children that has demonstrated the benefits on reading comprehension of training children to use mental imagery (Francey & Cain, 2015; Joffe et al., 2007). The inconsistent findings across studies potentially arise from diverse child and item characteristics. To date, there is a dearth of studies that have systematically investigated the associations between use of vocabulary scaffolds and young readers' comprehension of short digital text during independent reading, while accounting for both child and item characteristics. This is an innovative feature of our study.

Child Characteristics

Vocabulary scaffolds can be made available to the learner on an as-needed basis. This approach assumes that readers are aware of their vocabulary and comprehension support needs as they progress through the text. Such metacognitive awareness is likely to vary as a function of reader characteristics, such as ability level and age (Baker & Cerro, 2000; Kirby & Moore, 1987), gender (Sadeghi & Khezrlou, 2012; Tseng et al., 2006; Wu, 2014), and degree of bilingualism (Abu Rabia, 2019). To the best of our knowledge, research on individual differences in the selection of vocabulary scaffolds embedded in digital texts is lacking. However, reading skills predict 11-year-old students' selection of cohesive, semantically related hyperlinks (Salmerón & García, 2011), suggesting that literacy skills may influence how young readers go about navigating and searching for information in digital texts. With respect to gender, it has been shown that female students between 11 to 25 years employ more self-regulated strategies in vocabulary learning when reading texts compared to male students (Sadeghi & Khezrlou, 2012). Additionally, girls aged 5 to 7 years have been observed to outperform boys on a task designed to assess novel word learning, but only when learning phonologically or semantically familiar information (Kaushanskaya et al., 2013), suggesting that girls use different strategies compared to boys when learning novel words. Girls also outperform boys in knowledge of metacognitive strategies and navigation skills when engaged in online reading (Wu, 2014). Regarding language status, bilingualism is also associated with increased use of metacognitive reading strategies (Abu Rabia, 2019).

Child characteristics may also modulate the associations between use of vocabulary scaffolds and reading comprehension. This association may depend on both the child's background knowledge and their use of strategies. For example, meta-analytic evidence documenting the process of acquiring novel words in a foreign language suggests that the impact of scaffolds is greater for beginner learners than advanced learners (Yun, 2011). To date, little is known about which child characteristics predict the use of scaffolding features and modulate the associations between use of vocabulary scaffolds and reading comprehension. This information is essential to understand who can benefit most from this feature, and who may need additional support.

Item Characteristics

Another dimension to consider is that word class, as well as the inherent psycholinguistic properties of the scaffolded words, may pose different challenges to learners, and thus modulate the associations between use of vocabulary scaffolds and reading comprehension. Studies on early language acquisition suggest that novel nouns are learnt more easily than novel verbs (Childers & Tomasello, 2002; Gentner, 1982; Goldin-Meadow et al., 1976) and the same advantage for nouns is apparent in second language acquisition (Ellis & Beaton, 1993). Nevertheless, the concept of noun-bias has been challenged: Fourth graders demonstrate a learning advantage for non-nouns (i.e., verbs, adjective and adverbs) compared to nouns during story reading (Schwanenflugel et al., 1997). These mixed findings emphasize the need to take word class into account when examining the associations between use of vocabulary scaffolds and reading comprehension. Psycholinguistic properties of the words may also influence how easily they are learned. For instance, concreteness influences children's ability to learn words (Hadley et al., 2016). Moreover, the frequency of a word in child-directed speech and imageability ratings are both important predictors of lexical development (Hansen, 2017).

Present Study

The research evidence reviewed suggests that vocabulary scaffolds embedded in digital texts can play an important role in word learning. However, the strength of the association between use of scaffolds and reading comprehension remains unclear, with mixed findings arising potentially from diverse participant and item

characteristics. The current study was designed to address these gaps in our knowledge using secondary data analysis to answer the following critical research questions:

1) Do child characteristics predict the use of vocabulary scaffolds?

2) Is the use of vocabulary scaffolds associated with reading comprehension?

3) Is the association between use of scaffolds and reading comprehension conditional on child and/or item characteristics?

We examined the performance of a large cohort of young readers in the United States in a digital reading environment. During a gamified comprehension task, students acted as a newspaper editor and judged whether picture-text pairs matched in meaning or not. The task was designed to promote interest level, which has been shown to enhance comprehension monitoring in children with poor reading comprehension (De Sousa & Oakhill, 1996). Throughout the game, vocabulary scaffolds for target words were signalled by underlining. The student could select the word to get audio and visual support for its pronunciation and meaning. Our analysis examined the use of the scaffolds and the associations between use of scaffolds and reading comprehension as evidenced by performance on the sense-matching task, across a range of child and item characteristics.

Research Question 1: Do child characteristics predict the use of vocabulary scaffolds?

Our first research question investigated child characteristics as predictors of scaffold use. We examined four child characteristics: Literacy skill, grade (as a proxy for chronological age), gender, and language status (monolingual vs bilingual students). Whilst previous research on hyperlink selection strategies shows that reading ability influences how children search information and navigate digital texts (Salmerón & García, 2011), the reading environment and type of scaffold in that study differed substantially from our own, so we did not make directional predictions about the influence of reading skills and age on scaffold use. With respect to gender, however, we predicted that girls would be more likely to use scaffolds, because previous research has shown that girls are more likely to use self-regulated strategies in vocabulary learning (Sadeghi & Khezrlou, 2012), and have better knowledge of metacognitive strategies and navigation skills in online reading (Wu, 2014). Regarding language status, we hypothesized that bilingual students would make greater use of

scaffolds, since bilingualism is associated with an increased use of metacognitive reading strategies (Abu Rabia, 2019).

Research Question 2: Is the use of vocabulary scaffolds associated with reading comprehension?

Our second research question investigated the associations between use of vocabulary scaffolds and reading comprehension. We predicted that scaffold use would be associated with greater item accuracy (i.e., reading comprehension), given the strong relationship between vocabulary and reading comprehension (Cain et al., 2003) and previous meta-analytic evidence (Abraham, 2008). However, we note the findings of a recent meta-analysis indicating that an embedded dictionary aided word learning had a negligible or negative impact on story comprehension (Furenes et al., 2021).

Research Question 3: Is the association between use of scaffolds and reading comprehension conditional on child and/or item characteristics?

Our third research question addressed whether the associations between use of vocabulary scaffolds and reading comprehension was modulated by child and/or item characteristics. Our child characteristics of interest were early literacy skills, grade, gender, and language status. Yun's (2011) meta-analysis showed that the impact of similar scaffolds was greater in beginner as opposed to more advanced adult learners, during the acquisition of novel words in a foreign language. If those findings generalize to children, and to the domain of reading comprehension, the associations between scaffold use and reading comprehension should be positive and greater for children with lower literacy skills and beginner readers. Due to a lack of previous studies in this area examining the influence of gender and language status, we did not make specific predictions about how gender and language status might modulate the associations between use of vocabulary scaffolds and reading comprehension.

Lastly, we hypothesize that the associations between use of vocabulary scaffolds and reading comprehension would differ by word type, because they pose different challenges to learners (Childers & Tomasello, 2002; Ellis & Beaton, 1993; Gentner, 1982; Goldin-Meadow et al., 1976; Schwanenflugel et al., 1997). Specifically, we predicted that the associations between use of vocabulary scaffolds and reading comprehension would be positive and stronger for nouns than for other word types,

due to research demonstrating that nouns are easier to learn (Childers & Tomasello, 2002), and also because nouns are typically more concrete (Goldin-Meadow et al., 1976), making them easier to depict and define with the type of scaffold under investigation.

Methods

To address our research questions, we exploited an existing large observational dataset. The dataset was supplied through a project-specific data sharing agreement with *Amplify Education,* a private company dedicated to the development and commercialization of digital educational tools. The dataset consists of behavioural/usage data (i.e., user log files) relative to the *Reveal Words* vocabulary scaffolding feature embedded in the *Read All About It* reading game. The data procured was for students who played the game between 01/09/2020 and 28/04/21, as part of a broader digital reading supplement, Amplify Reading. This study has received Ethical approval from the Faculty of Science and Technology, Lancaster University (reference number: FST-2022-1098-RECR-5). Our analysis plan was preregistered and is available on the Open Science Framework (OSF): Diprossimo, L., Cain, K., & Ushakova, A. (2021, June 9). Vocabulary Scaffolding Features and Young Readers' Comprehension of Digital Text: Insights from a Big Observational Dataset. https://doi.org/10.17605/OSF.IO/62C4Q.

Participants

Children across the U.S., primarily based in Southern California and Texas, played the *Read All About It* game as part of a broader digital reading supplement. Observations with missing data on key variables of interest (i.e., literacy skills, grade level[3], gender[4], and language status) were removed. We then randomly sampled within the large observational dataset to reduce it to half of all cases per each game item/use of scaffold combination. This was done to save computational time while preserving the dataset's inherent characteristics. With large enough samples, there is a certain point after which increasing the sample size will not affect the results but may

[3] *Grade level* was our proxy for chronological age, as precise information on chronological age was not available in the anonymised dataset.

[4] Here we use the term *gender* in line with the previous literature and to acknowledge the fact that literacy is embedded in a socio-cultural context. Strictly speaking however, we classed our participants on reported biological sex.

bring additional computational challenges. After subsampling, our sample was still much larger (N ~120,000) than those commonly reported in the literature. English monolingual and bilingual students[5] aged between 5 and 8 years were included. The data set used for this analysis included only student data from classroom use of *Amplify Reading*. No personally identifiable information (for either children, schools, or school districts) was included in the dataset that was analysed.

Early literacy skills were measured for each child with the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), 8th edition, (University of Oregon, Center on Teaching and Learning, 2018). DIBELS is a set of short, standardized assessments that can be used for universal screening, benchmark assessment, and progress monitoring. DIBELS was administered by school teachers to evaluate letter naming fluency, phonemic segmentation, nonsense word fluency, word reading fluency, oral reading fluency, and reading comprehension (using a maze – or cloze – task). The scores of the different DIBELS subtests enabled the classification of students into four composite performance levels: *well below benchmark*, *below benchmark*, *at benchmark*, and *above benchmark*. Our analysis employed composite performance levels as an overall measure of literacy skills. The composition of our sample is summarized in Table 14.

[5] Here we use the term *bilingual* in an inclusive way, encompassing a variety of types of bilingualism. Our dataset does not permit distinction between, for example, simultaneous or sequential bilingualism. Examination of these more fine-grained distinctions were, therefore, beyond the scope of this study.

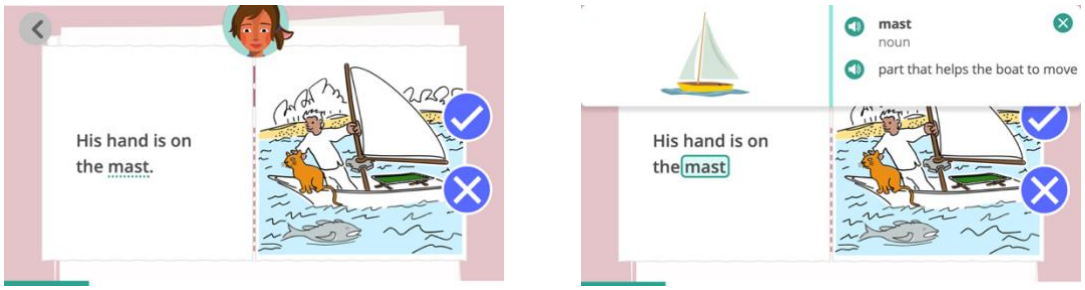
Table 14 *Descriptives for the Key Variables of Interest in the Study Sample: Literacy Skills, Grade, Gender, and Language Status. Information on Reported Race is also Provided*

		Count	Proportion (%)
Literacy skills (DIBELS)	Above Benchmark	26956	23.1
	At Benchmark	29024	24.87
	Below Benchmark	18163	15.56
	Well Below Benchmark	42570	36.47
Grade	K	21369	18.31
	1	32673	27.99
	2	41775	35.79
	3	20896	17.9
Gender	Male student	59019	50.57
	Female student	57694	49.43
Language status	English monolingual	90636	77.66
	Bilingual student	26077	22.34
Race	Black or African American	18471	15.83
	White	22813	19.55
	Hispanic or Latino	61884	53.02
	Asian	5256	4.5
	Multiracial/other	2387	2.05
	Native Hawaiian, Pacific Islander, American Indian, Alaskan Native	2386	2.04
	Not available	3516	3.01

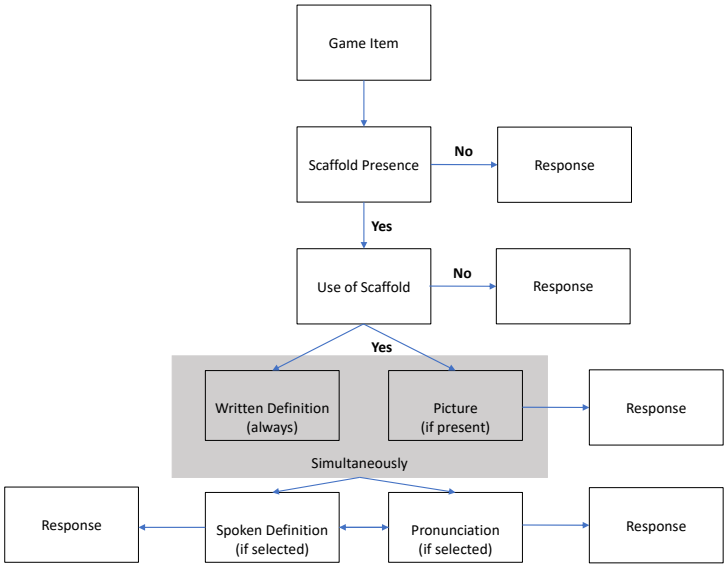
Learning Materials

Read All About It is a reading game designed to foster automaticity in reading. In a gamified comprehension task, students are presented with text-picture pairs that assess understanding of a target item. They are encouraged to act as newspaper editors and judge whether the text matches the picture by selecting (i.e., mouse clicking or tapping) either the correct (i.e., V) or incorrect (i.e., X) option (Figure 4a). The decodable text includes the sound-spelling correspondences, word features (e.g., prefixes/suffixes), and phonics rules that they have learned and practised in other games. The game is introduced after students have been taught 10 individual letter sounds. Content builds as students learn new letter-sound correspondences, letter combinations, and phonics rules through multisyllabic decoding. Content difficulty increases as the child progresses to more advanced levels. Each child is initially assigned to a game level based on their reading ability, and provisionally removed from a game level if they repeatedly fail; this is done to minimize their frustration.

Figure 4 *Example of Game Item, Scaffolding Feature and Child's Decision Tree*



a) b)



c)

Note. (a) Gamified reading comprehension item, (b), *Reveal Word* scaffolding feature for the word "mast", (c) child's decision tree while navigating an item in the game.

Throughout the game, the *Reveal Words* scaffold (Figure 4b) appears with the target word underlined. The student may select this word to get immediate support for its meaning via written definition and pictorial support, which appears simultaneously. We note however that pictorial support was not provided for a small proportion of scaffolded words[6]. Auditory support for pronunciation and definition were available by further clicking on the speaker symbols (see Figure 4c for a sketch). *Reveal Words* scaffolds were available for a range of word types (i.e., nouns, adjective,

[6] Overall, pictorial support was available for 71% of the scaffolds. We return to this point in our exploratory analysis.

adverbs, verbs, and prepositions)[7]. The game comprises a total of 54 game levels, 505 game items, and 581 *Reveal Words* scaffolds. The type and number of *Reveal Words* scaffolds varied across game items. The words that received *Reveal Words* support were selected based on the Biemiller ratings (Biemiller, 2010), playtesting feedback, and the English Vocabulary Profile (www.englishprofile.org). The full list of scaffolded words is available in the Supplementary Materials.

Procedures

Children played the *Read All About It* game via a tablet or a PC in school settings. The data collection was a continuous process, as data were collected iteratively, every time a child logged into the game. Log files provided information about usage, such as *Reveal Words* feature selection, and accuracy for the comprehension task. Each log file was associated with a unique identification code for each child. Early literacy skills, grade, gender, and language status were also encoded in the log files. A codebook with full list of information available in the log files can be consulted in the Supplementary Materials.

Analytical Approach

Statistical Models

Research Question 1: Do child characteristics predict the use of vocabulary scaffolds? To investigate which child characteristics influenced the probability of using at least one scaffold per game item, we fitted a Generalized Linear Mixed Model (GLMM; Baayen 2008) with binomial error structure and logit link function (McCullagh & Nelder, 1989). Our binary outcome was the use of the scaffold. We included early literacy skills as measured by DIBELS, grade level (as a proxy for age), gender, and language status, all as fixed effects; we further included game item and level as random effects to account for potential differences arising from individual game items and levels.

Research Question 2: Is the use of vocabulary scaffolds associated with reading comprehension? Similar to the first analysis, we fitted a Generalized Linear Mixed Model (GLMM; Baayen 2008) with binomial error structure and logit link

[7] The scaffold was available for 311 nouns, 59 adjectives, 20 adverbs, 188 verbs, and 2 prepositions. Pictorial support was provided for 77% of nouns, 63% of adjective, 25% of adverbs, 69% verbs and 100% of prepositions.

function (McCullagh & Nelder, 1989) to investigate what influenced the probability of being correct in each game item. Here our binary outcome was the accuracy in each game item. We included the use of at least one scaffold per game item as a test predictor, and literacy skills as measured by DIBELS, grade, gender, and language status, as control variables. As before, we controlled for the random effect of each game item and level to account for potential differences arising from individual game items and levels.

Research Question 3: Is the association between use of scaffolds and reading comprehension conditional on child and/or item characteristics? We further expanded our previous model by including the interactions between the use of the scaffold and each of the child characteristics of interest (i.e., early literacy skills, grade, gender, and language status) as well as the number of scaffolded nouns, adjectives, adverbs, verbs, and prepositions as test predictors. As before, we controlled for the random effect of each game item and level to account for potential differences arising from individual game items and levels.

Model Selection

Model selection was guided by the results of the incremental likelihood ratio test (Matuschek et al., 2017), which indicates whether a specific term significantly improves the model fit. The selection process started at the baseline model (a model lacking our test predictor), adding a single predictor at a time until we reached the full model, as indicated in our pre-registration (https://doi.org/10.17605/OSF.IO/62C4Q).

Inference Criteria

To evaluate the significance of the contribution of our predictors, whilst avoiding multiple testing (Forstmeier & Schielzeth, 2011), we compared our full models with null models lacking our test predictors but being otherwise identical using the likelihood ratio test. The significance of the beta coefficients was indicated by $p < .05$. Null hypothesis significance testing was complemented by examination of odds ratios and relative confidence intervals (see tables in Supplementary Materials). Odds ratio provided an index of the strength of the relationships between our predictors and outcome, thus enabling meaningful comparison between predictors. Marginal effects were plotted to provide straightforward visualization of predicted probabilities for the results of the more complex models (Lüdecke, 2018). Significant interaction terms

accompanied by improvement in the model fit, as indicated by a significant likelihood ratio test, indicated the presence of conditional effects. Marginal R^2 indicated the variance explained by the fixed effects, while conditional R^2 indicated the variance explained by both the fixed and random effects.

Model Implementation

The models were implemented in R (version 4.0.4. 2021-02-15) with the function glmer of the R package lme4 (version 1.1- 26) (Bates et al., 2015). Marginal effects (i.e., predicted probabilities) were computed using the function ggpredict of the R package ggeffects (Lüdecke, 2018). Predicted probabilities plots are provided for visualisation purposes and were used to guide the interpretation of interaction terms. Best practice guidelines for the reporting of mixed models (Meteyard & Davies, 2020) were followed (i.e., we report the version of the software and packages used, data preparation steps, model selection and model output including coefficients, standard errors, confidence intervals and p-values, random effects, and R^2 as an index of model fit).

Exploratory Analysis

Exploring the Contribution of Pictorial Support, Pronunciation Support, Spoken and Written Definition and their Combinations.

We ran an exploratory analysis to examine the unique contributions of the different components of the scaffold (i.e., picture, pronunciation, written and spoken definitions) and their combinations (i.e., picture-pronunciation, picture-definition, pronunciation-definition, and all those together) on reading comprehension performance. To enable this, we subset our dataset and focused on instances where a single *Reveal Words* scaffold was used in each game item. This approach enabled us to quantify the proportion of correct responses across different components of the scaffold that were available and used, while fixing the number of scaffolds constant at 1 to ensure comparability across game items.

Do Age of Acquisition and Concreteness of Scaffolded Words Influence the Associations between Use of Scaffolds and Reading Comprehension?

We modelled an observational dataset, which was not designed for research purposes. Therefore, we checked whether key psycholinguistic variables such as age of acquisition and concreteness varied systematically across word types in our dataset.

We further explored whether age of acquisition and concreteness influenced the effectiveness of the scaffold. We analysed the same subset of data as that used in the previous exploratory analysis. By working with instances where only a single *Reveal Words* scaffold was used in each game item, we were able to examine the relation between the psycholinguistic properties of a specific scaffolded word and reading comprehension performance (indicated by game item accuracy). The age of acquisition and concreteness ratings used in this analysis were derived from open sources databases (Brysbaert et al., 2014; Kuperman et al., 2012).

Results

Confirmatory analysis

Research Question 1: Do Child Characteristics Predict the Use of Vocabulary Scaffolds?

A clear impact of early literacy skills, grade, gender, and language status on the probability of using the scaffolding feature was evident. This was indicated by the significant likelihood ratio test comparing the full model including our test predictors and the null model lacking our predictors but being otherwise identical (χ^2 = 21877.00, *df* = 8, $p < .001$). We discuss the output of the model (Table 15) first, followed by the model-generated predicted probability plots (Figure 5).

Table 15 *Results of the Model Estimating the Effects of Literacy Skills, Grade, Gender and Language Status on the Probability of Using the Reveal Word Scaffold*

	Use of the Scaffold			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	-0.76	0.06	-0.88 – -0.64	**<0.001**
DIBELS [Well Below Benchmark]	0.24	0.01	0.23 – 0.25	**<0.001**
DIBELS [Below Benchmark]	0.22	0.01	0.21 – 0.24	**<0.001**
DIBELS [Above Benchmark]	-0.31	0.01	-0.32 – -0.30	**<0.001**
Grade [1]	-0.56	0.01	-0.57 – -0.54	**<0.001**
Grade [2]	-1.21	0.01	-1.23 – -1.19	**<0.001**
Grade [3]	-1.44	0.01	-1.46 – -1.42	**<0.001**
Gender [Male]	-0.20	0.00	-0.21 – -0.19	**<0.001**
Language Status [Bilingual]	0.24	0.01	0.23 – 0.25	**<0.001**
Random Effects				
σ^2	3.29			
$\tau_{00\ q\_id}$	0.12			
$\tau_{00\ level\_name}$	0.37			
ICC	0.13			
$N_{\ q\_id}$	505			
$N_{\ level\_name}$	54			
Observations	1716787			
Marginal R^2 / Conditional R^2	0.065 / 0.186			

Note. Literacy skills [DIBELS] are dummy coded with 'at benchmark' set as the reference category. Grade is dummy coded with kindergarten set as the reference category. Language status is dummy coded with English monolingual set as the reference category. Gender is dummy coded with girls set as the reference. Model fit is indicated by R^2.

With respect to literacy skills, children classified as being *well below benchmark* and *below benchmark* were more likely to use the scaffolding feature compared to children classified as being *at benchmark*. This was indicated by the positive sign of the respective coefficients (β = 0.24, $p < .001$; β = 0.22, $p < .001$). In contrast, children classified as being *above benchmark* were less likely to use the scaffolding feature compared to those *at benchmark*. This was indicated by the negative sign of the coefficient (β = -0.31, $p < .001$). These results suggest that literacy skills predict the use of the scaffold. Students who need more support seek more support, as shown by the increased probability of using the scaffold in lower ability levels. The opposite was true for children in higher ability levels.

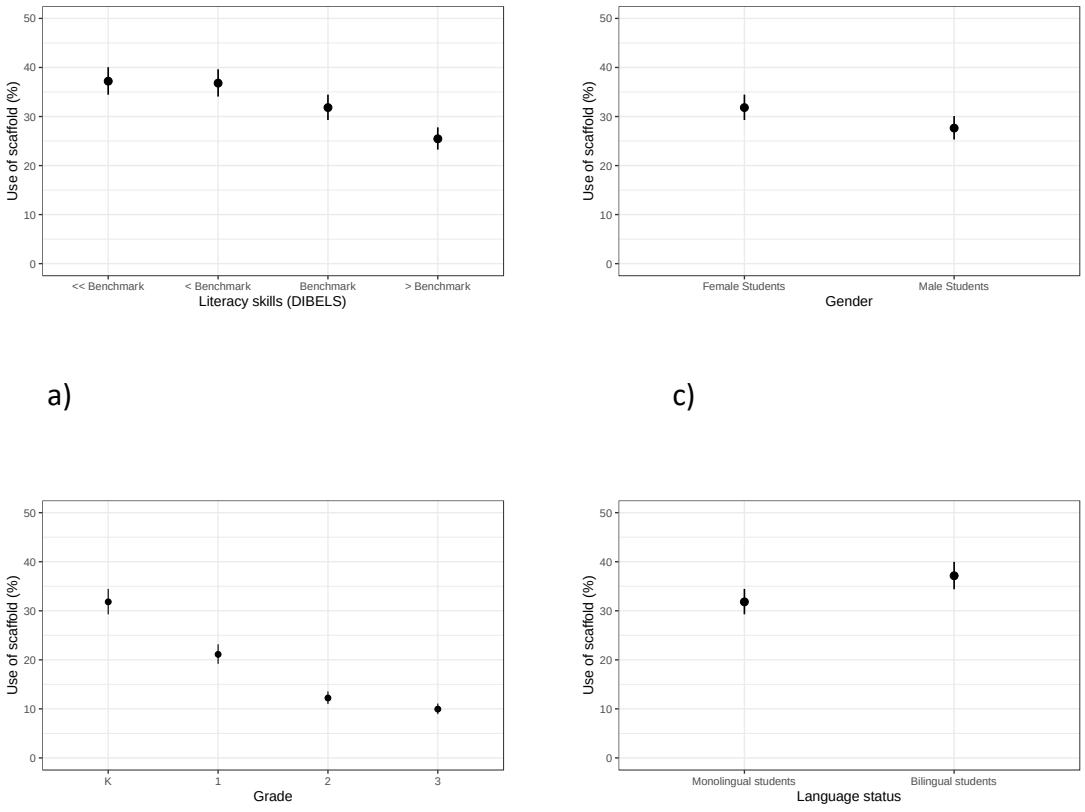
Grade served as a proxy for chronological age. A decreased probability of using the scaffold was associated with each of the higher grades, as compared to kindergarten. This was indicated by the negative sign of the coefficients for grades 1 (β = -0.56, $p < .001$), 2 (β = -1.21, $p < .001$), and 3 (β = -1.44, $p < .001$). This pattern of results suggests a progressive *decrease* in the probability of using the scaffolding feature as students move up through the grades (i.e., increase in age). Beginner readers were more likely to make use of the scaffold, and this probability decreased in more experienced readers.

We also found an influence of gender and language status. Boys were less likely to use the scaffold compared to girls, as indicated by the negative sign of the coefficient (β = - 0.20, $p < .001$). Furthermore, students reporting as bilingual were more likely to use the scaffold compared to those reporting as English monolingual, as shown by the positive coefficient (β = 0.23, $p < .001$).

The predicted probability of using the scaffold expressed in percentage on the y axis, with child characteristics on the x axis is reported in Figure 5. Each panel displays one of the four child characteristics of interest. Error bars represent the confidence intervals. Figure 5a shows the extent to which the predicted probability of using the scaffold varied across literacy skills. Children *well below benchmark* had the highest predicted probability of using the scaffold (~ 35%) and this probability was less (~ 25%) for children *above benchmark*. Figure 5b illustrates how the predicted probability of using the scaffold varied across grades. Children in kindergarten had the highest predicted probability of use (~ 30%), while children in grade 3 had the lowest (~ 10%).

Differences in the predicted probability of using the scaffold by gender are reported in Figure 5c. The predicted probability of using the scaffold for girls was above 30%, while for boys it was below 30%. Differences in the predicted probability of using the scaffold between monolingual and bilingual students are illustrated in Figure 5d. The predicted probability of using the scaffold for monolingual students was below 35%, while for bilingual students it was above 35%.

Figure 5 *Predicted Probability of Using the Scaffold across Child Characteristics*



Note. Predicted probability of using the scaffold by: (a) early literacy skills, (b) grade, (c) gender, and (d) language status, computed holding non-focal terms constant at their reference level. Error bars represent confidence intervals.

The strength of the relationships between child characteristics and use of scaffold as indicated by odds ratios is reported in the Supplementary Materials. Overall, the associations were modest, with odds ranging from 0.24 to 1.27. Children classified as being well below benchmark according to their literacy skills were 1.27 times more likely to use the scaffold as compared to children at benchmark.

Research Question 2: Is the Use of Vocabulary Scaffolds Associated with Reading Comprehension?

We examined whether use of the scaffold was associated with reading comprehension performance. Overall, the use of the scaffolding feature was associated with an increased probability of getting the associated item correct in the reading comprehension (text-picture matching) task. This was indicated by the significant likelihood ratio test comparing our full model with a null model lacking our test predictor but being otherwise identical (χ^2 = 2226.80, *df* = 1, $p < 0.001$). The results of the full model are reported in Table 16.

Table 16 *Results of the Model Estimating the Effects of the Scaffold, Literacy Skills, Grade, Gender and Language Status on Task Accuracy*

	Accuracy in the Task			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	1.34	0.04	1.26 – 1.42	**<0.001**
Use of Scaffold [yes]	0.26	0.01	0.25 – 0.27	**<0.001**
DIBELS [Well Below Benchmark]	-0.39	0.01	-0.40 – -0.38	**<0.001**
DIBELS [Below Benchmark]	-0.10	0.01	-0.11 – -0.09	**<0.001**
DIBELS [Above Benchmark]	0.23	0.01	0.22 – 0.24	**<0.001**
Grade [1]	-0.01	0.01	-0.03 – 0.00	0.050
Grade [2]	-0.03	0.01	-0.05 – -0.02	**<0.001**
Grade [3]	0.16	0.01	0.14 – 0.18	**<0.001**
Gender [Male]	-0.09	0.00	-0.10 – -0.08	**<0.001**
Language Status [Bilingual]	-0.21	0.00	-0.22 – -0.21	**<0.001**
Random Effects				
σ^2	3.29			
$\tau_{00\ q\_id}$	0.70			
$\tau_{00\ level\_name}$	0.04			
ICC	0.18			
$N_{\ q\_id}$	505			
$N_{\ level\_name}$	54			
Observations	1716787			
Marginal R^2 / Conditional R^2	0.020 / 0.200			

Note. Literacy skills [DIBELS] are dummy coded with at benchmark set as the reference category. Grade is dummy coded with kindergarten set as the reference category. Language status is dummy coded with English monolingual set as the reference category. Gender is dummy coded with girls set as the reference. Model fit is indicated by R^2.

Use of the scaffolding feature was associated with an increased probability of getting the item correct compared to not using the scaffold. This is indicated by the positive sign of the coefficient (β = 0.26, $p < .001$). Importantly, the effect of the scaffold was significant after controlling for child literacy skills, grade, gender, and language status. Of note, each of the four control predictors was significant and indicated that children with lower literacy skills, boys, and bilingual students were less likely to get the item correct in the reading comprehension task. In addition, children in grade 2 were less likely to get the item correct compared to kindergarteners while children in grade 3 were more likely to get the item correct compared to kindergarteners.

Odds ratios are reported in Supplementary Materials to illustrate the strength of the association between use of scaffold and reading comprehension. This association was modest (*OR* = 1.30). We consider the practical significance of these findings in the discussion.

Research Question 3: Is the Association between Use of Scaffolds and Reading Comprehension Conditional on Child and/or Item Characteristics?

In these analyses, we further investigated the interactions between the use of the scaffold and child characteristics and the influence of item characteristics (word type) on the probability of getting the game item correct. The comparison of the full and null model (χ^2 = 1350.00, *df* = 13, $p < .001$) revealed significant interactions between use of scaffold and child characteristics and a significant effect of word type. The results of the full model are reported in Table 17.

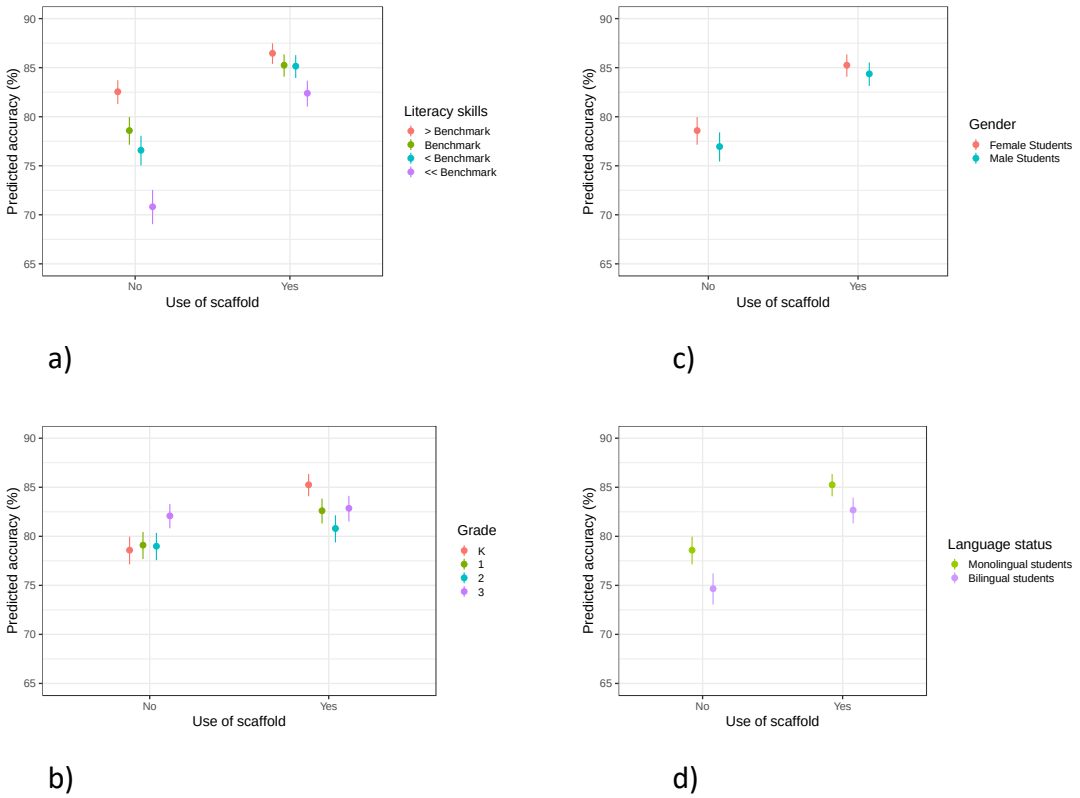
Table 17 *Results of the Model Estimating the Effects of the Scaffolded Word Type, and Interaction Between Scaffold Use and Literacy Skills, Grade, Gender, and Language Status on the Accuracy in the Task*

	Accuracy in the Task			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	1.30	0.04	1.22 – 1.38	**<0.001**
Use of Scaffold [yes]	0.45	0.02	0.41 – 0.49	**<0.001**
DIBELS [Well Below Benchmark]	-0.41	0.01	-0.42 – -0.40	**<0.001**
DIBELS [Below Benchmark]	-0.12	0.01	-0.13 – -0.10	**<0.001**
DIBELS [Above Benchmark]	0.25	0.01	0.24 – 0.26	**<0.001**
Grade [1]	0.03	0.01	0.02 – 0.05	**<0.001**
Grade [2]	0.02	0.01	0.01 – 0.04	**0.009**
Grade [3]	0.22	0.01	0.20 – 0.24	**<0.001**
Gender [Male]	-0.09	0.00	-0.10 – -0.09	**<0.001**
Language Status [Bilingual]	-0.22	0.00	-0.23 – -0.21	**<0.001**
Use of Scaffold [yes]:DIBELS [Well Below Benchmark]	0.20	0.01	0.17 – 0.23	**<0.001**
Use of Scaffold [yes]:DIBELS [Below Benchmark]	0.11	0.02	0.08 – 0.14	**<0.001**
Use of Scaffold [yes]:DIBELS [Above Benchmark]	-0.15	0.01	-0.18 – -0.13	**<0.001**
Use of Scaffold [yes]:Grade [1]	-0.23	0.02	-0.26 – -0.20	**<0.001**
Use of Scaffold [yes]:Grade [2]	-0.34	0.02	-0.38 – -0.31	**<0.001**
Use of Scaffold [yes]:Grade [3]	-0.40	0.02	-0.44 – -0.36	**<0.001**

Use of Scaffold [yes]:Gender [Male]	0.03	0.01	0.01 – 0.05	**0.015**
Use of Scaffold [yes]:Language Status [Bilingual]	0.03	0.01	0.00 – 0.05	**0.023**
Scaffolded Nouns	0.03	0.01	0.01 – 0.05	**0.001**
Scaffolded Verbs	-0.02	0.01	-0.04 – 0.00	0.057
Scaffolded Adjectives	-0.35	0.02	-0.40 – -0.30	**<0.001**
Scaffolded Adverbs	-0.18	0.05	-0.28 – -0.07	**0.001**
Scaffolded Prepositions	-0.02	0.06	-0.14 – 0.10	0.712
Random Effects				
σ^2	3.29			
$\tau_{00\ q\_id}$	0.69			
$\tau_{00\ level\_name}$	0.05			
ICC	0.18			
$N_{\ q\_id}$	505			
$N_{\ level\_name}$	54			
Observations	1716787			
Marginal R^2 / Conditional R^2	0.021 / 0.201			

Note. Literacy skills [DIBELS] are dummy coded with at benchmark set as the reference category. Grade is dummy coded with kindergarten set as the reference category. Language status is dummy coded with English monolingual set as the reference category. Gender is dummy coded with girls set as the reference. Model fit is indicated by R^2.

Figure 6 *Predicted Probability of Accuracy in the Task by Use of Scaffold across Child Characteristics*



Note. Predicted probability of accuracy (y axis) plotted against the use of scaffold (x axis) across child characteristics (colour coded) for: (a) literacy skills, (b) grade, (c) gender, and (d) language status. Predicted probability are computed holding non-focal terms constant at their reference level.

To guide interpretation of the interaction between scaffold use and child characteristics, we provide visualization of predicted probabilities (Figure 6). The significant interactions between scaffold use and literacy skills (β = 0.20, p < .001; β = 0.11, p < .001; β = -0.15, p < .001) suggest that the association between use of scaffolds and reading comprehension was modulated by literacy skills. As illustrated in Figure 6a, use of the scaffold was associated with higher performance for all children, but to a different extent depending on their literacy skills. Specifically, this association was greater for readers at lower ability levels. In other words, the use of the scaffold tended to minimise performance differences between ability levels. The significant

interactions between scaffold use and grade (β = -0.23, p < .001; β = -0.34, p < .001; β = -0.40, p < .001) suggest that the association between use of scaffolds and reading comprehension was modulated by grade. Specifically, the predicted probability plot (Figure 6b) illustrates that, although the use of the scaffold was associated with higher performance for all children, this association was greater for younger children such as those in kindergarten.

There was a significant interaction between scaffold use and gender (β = 0.03, p = .01). The predicted probability plot (Figure 6c) illustrates that the differences in performance due to gender were slightly reduced when the scaffold was used. Finally, there was a significant interaction between scaffold use and language status (β = 0.03, p = .02). The predicted probability plot (Figure 6d) shows that performance differences associated with language status were reduced when the scaffold was used.

The number of scaffolds used for nouns in each game item had a positive impact on the probability of getting the game item correct, as compared to not using a scaffold for nouns, as indicated by the positive sign of the coefficient (β = 0.03, p < .001). That is, for nouns, use of the scaffold was associated with better performance. In contrast, the number of scaffolds used for adjectives and adverbs in each game item had a negative impact on accuracy, as compared to not using a scaffold for adjectives and adverbs respectively. This is indicated by the negative sign of the respective coefficients (β = -0.35, p < .001; β = -0.18, p = .001). There was no clear effect of the number of scaffolds used for verbs and prepositions on item accuracy (β = -0.02, p = .06; β = -0.02, p = .71). Odds ratios are reported in the Supplementary Materials to illustrate the strength of the associations under investigations. Similar to previous analyses, the associations were modest.

Exploratory Analysis

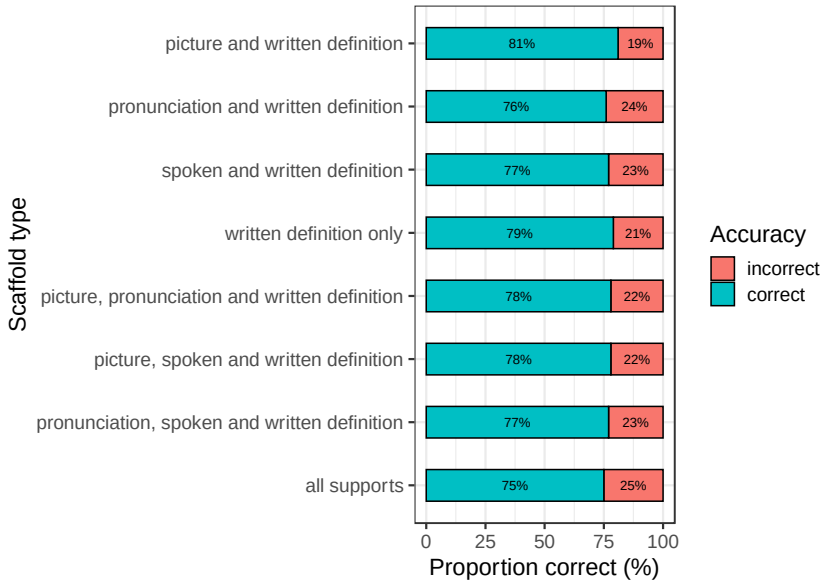
Exploring the Contribution of Pictorial Support, Pronunciation Support, Spoken and Written Definition and their Combinations

We conducted an exploratory analysis with the aim to disentangle the associations between use of different components of the scaffold (i.e., picture, pronunciation, written and spoken definitions) and their combinations (i.e., picture-pronunciation, picture-definition, pronunciation-definition, and all those together) and

reading comprehension performance. Written definitions were available for all *Reveal Words* scaffolds, whilst pictures were available for most, but not all (see learning materials in the methods section). Of note, when children clicked on the scaffold, written definitions and pictures (if available) were immediately visible, whereas children had to click further on the speaker icon to obtain the pronunciation and spoken definition supports. To examine the contributions of these distinct scaffold components and their combinations, we subset our dataset and focused on instances where a single *Reveal Words* scaffold was used in each game item. This was because some items had scaffolds for more than a single word. This approach enabled us to quantify how the proportion of correct responses varied across the components of the scaffold that were available and used, while fixing the number of scaffolds constant at 1 to ensure comparability across game items. Descriptive statistics revealed that the proportion of correct responses was broadly comparable across the different scaffold components and their combinations, with the proportion of correct responses ranging between 75 to 81% (Figure 7).

These descriptives also suggest that the effects of different components of a scaffold were not additive. For instance, the proportion of correct responses when all the supports were available and used (i.e., picture, pronunciation, spoken and written definition) was lower (75%) compared to when only the picture and written definition support were available and used (81%), or only the written definition was available (79%). Overall, pictorial support and written definition in combination were associated with the highest proportion of correct responses, followed by written definition alone (proportion correct 81% and 79%, respectively).

Figure 7 *Proportion of Correct Responses across Scaffold Components and Combinations*

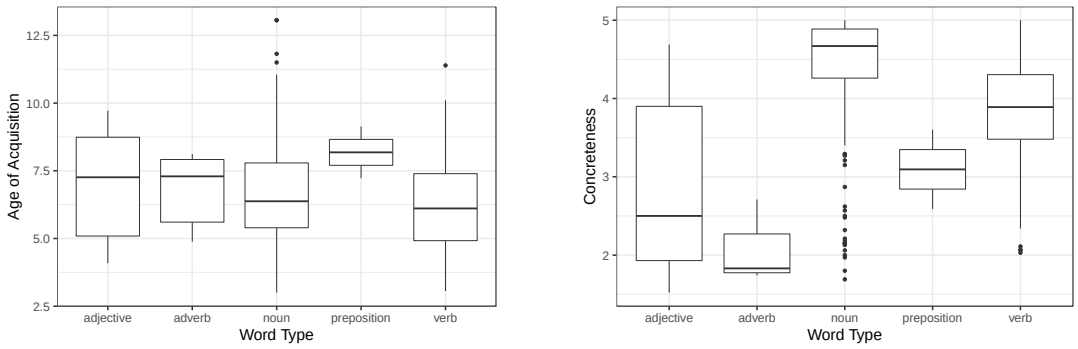


Note. This analysis was conducted on a subset dataset where only a single scaffold per game item was used to enable meaningful comparisons.

Do Age of Acquisition and Concreteness of Scaffolded Words Modulate the Associations between Use of Scaffolds and Reading Comprehension?

Because we modelled an observational dataset we checked whether key psycholinguistic variables of the words, such as age of acquisition and concreteness, varied systematically across word types in our dataset. To this end, we produced the boxplots illustrated in Figure 8. The horizontal line in the body of each boxplot represents the median value of age of acquisition for each category of word. The body of the boxplot represents the interquartile range. The vertical line in each boxplot signals the minimum and maximum value. Dots represent potential outliers. Figure 8 illustrates that age of acquisition was comparable across word type (Figure 8a), with median age of acquisition between 6 and 8 years of age (i.e., the age range of our participating children). In contrast, median concreteness differed across word types (Figure 8b) and may thus act as a confounding variable. We will return to this point in the discussion.

Figure 8 *Age of Acquisition and Concreteness across Word Types*

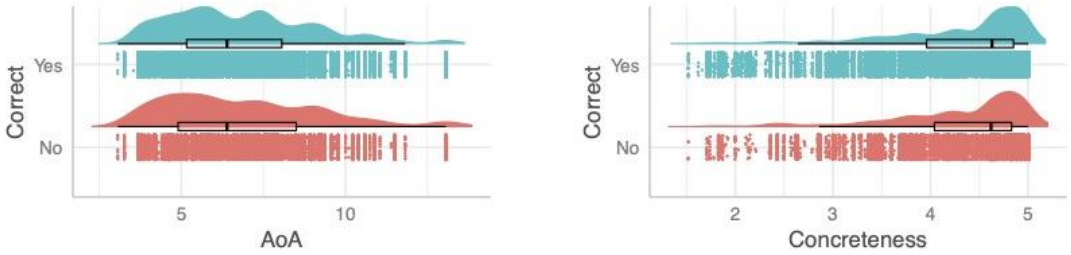


a) b)

Note. Box plots of (a) age of acquisition, and (b) concreteness, across word type

To explore whether age of acquisition and concreteness of scaffolded words influence the associations between use of scaffolds and reading comprehension, we analysed the same subset of data used in the previous exploratory analysis. As before, we worked with instances where only a single *Reveal Words* scaffold was used in each game item. This enabled us to examine the relation between the psycholinguistic properties of a specific scaffolded word and reading comprehension performance as evidenced by game item accuracy. The results are illustrated with raincloud plots in Figure 8. Correct and incorrect responses are shown on the y axis and age of acquisition (Figure 8a) and concreteness (Figure 8b) on the x axis. These visualizations combine boxplots (see previous paragraph for interpretation), probability density, and the jittered raw dataset (M. Allen et al., 2019). This comprehensive data overview suggests that, in a scenario where children use a single scaffold, the psycholinguistic properties of the scaffolded word do not influence accuracy. Indeed, the median and distribution of age of acquisition and concreteness of *Reveal Words* scaffolds is comparable across correct and incorrect responses. This pattern suggests that the age of acquisition and concreteness of scaffolded words do not influence the associations between use of scaffolds and reading comprehension.

Figure 9 *Age of Acquisition and Concreteness across Correct and Incorrect Responses*



a) b)

Note. Raincloud plots showing (a) age of acquisition and (b) concreteness on the x axis, across correct and incorrect responses on the y axis, in a scenario where a single scaffold was used.

Discussion

Our findings advance our understanding of vocabulary scaffolding features embedded in digital reading environments in several ways. First, we found that child characteristics predicted the use of scaffolds: Children with lower literacy skills, beginner readers, girls, and students who reported as bilingual were more likely to use the scaffold. Second, the use of the scaffold was associated with better reading comprehension. This association was stronger for three of the four child characteristics associated with greater scaffold use: Children with lower literacy skills, beginner readers, and bilingual students. In addition, this association was stronger for boys than for girls. Third, the association between the use of scaffold and reading comprehension varied across word types, with a positive association for nouns, a negative association for adjective and adverbs, and no evidence of an association for verbs and prepositions. We discuss each finding in relation to the previous literature, followed by recommendations for future research and educational implications.

Child Characteristics Predicted the Use of Scaffolds

Our first research question sought to establish which child characteristics predicted the use of vocabulary scaffolds that were available to learners on an as-needed basis. Children with lower literacy skills and beginner readers were more likely to use the scaffold. This suggests that struggling and beginner readers have sufficient

metacognitive awareness to seek support while navigating a digital reading environment. In addition, girls and bilingual students were more likely to use the scaffold. These findings are in line with evidence for enhanced metacognitive skills in girls (Wu, 2014) and bilingual readers (Abu Rabia, 2019). On the one hand, these results are encouraging as they show that children with lower literacy skills, beginner readers, and bilingual readers, all groups that are associated with lower reading comprehension (Catts et al., 2006; LARRC, 2015; Lesaux & Kieffer, 2010), are more likely to use a scaffold. On the other hand, these results point to the need to further prompt the use of such scaffolds amongst male students. Our analyses showed that boys were less accurate compared to girls, and that performance differences due to gender were reduced when the scaffold was used. This has important and positive implications for educational practice, due to the common finding of lower reading attainment in boys than in girls (Duncan et al., 2016; Wu, 2014). In addition, we note that it may be beneficial to prompt the use of the scaffolding feature across all students, as the overall use of the scaffold was lower than desirable, and scaffold use was associated greater reading comprehension. Two potential ways to promote the use of the scaffold are the addition of extended on-boarding instructions reviewing how to use the feature, and/or initial adult modelling. These strategies have the potential to promote children's independent use of the scaffold. App developers should also consider providing these types of scaffolds based on a student's performance, in an adaptive manner (Sampayo-Vargas et al., 2013).

The Use of Scaffolds was Associated with Better Reading Comprehension

Critically, we found that the use of the vocabulary scaffolding feature was associated with better reading comprehension of short digital texts. These results extend previous meta-analytic evidence reporting a positive effect of similar scaffolds on the reading comprehension of university students when reading in a foreign language (Abraham, 2008). The benefit of vocabulary scaffolds appears to generalize to younger students, such as those included in our study (i.e., 5 to 8 years), and to first language reading comprehension[8]. Our findings also add to the literature demonstrating that other types of scaffold, such as graphic organizers, can benefit

[8] Our sample included both English monolingual and bilingual students, both reading in English, thus providing evidence for first language reading comprehension.

reading comprehension (Elbro & Buch-Iversen, 2013). However, our results contrast with meta-analytic evidence showing a null or negative effect on children's story comprehension when dictionary supports are embedded in storybooks (Furenes et al., 2021). These inconsistencies are likely to arise from the longer texts used in the storybooks examined in other research, and possibly from the co-reading context that was the focus of the meta-analysis. Future research should determine any specific influence of the context (educational, recreational, individual, co-reading), as well as length of text, to inform specific recommendations for the inclusion of this support.

The Association between Use of Scaffolds and Reading Comprehension Was Modulated by Child and Item Characteristics

Our study provides the first evidence that the association between use of scaffolds and reading comprehension was modulated by both child and item characteristics. Specifically, our analysis suggested that the associations between the use of scaffolds and reading comprehension was greater for children with lower literacy skills and beginner readers. This extends previous research showing that the beneficial effects of diagrams as scaffolds for learners' cognitive and metacognitive processes is stronger for participants with low verbal ability (Cuevas et al., 2002). Similarly, our findings align with a meta-analysis suggesting that the impact of similar scaffolds is greater in beginner learners when acquiring novel words in a foreign language (Yun, 2011) and extends those findings to the domain of reading comprehension.

Additionally, we found that performance differences due to gender and language status were reduced when the scaffold was used. Because boys and bilingual students are typically found to have poorer reading comprehension skills (Duncan et al., 2016; Wu, 2014), our findings are especially encouraging. In short, the use of scaffolds is associated with higher performance especially in typically lower achieving student subgroups. This has important educational implications and speaks in favour of embedding scaffolding features in digital texts, especially for children with or at risk of reading comprehension difficulties.

Regarding item characteristics, the use of scaffolds was associated with higher performance for nouns and lower performance for adjective and adverbs. No differences emerged for verbs and prepositions. The advantage observed for

scaffolded nouns can be explained in two ways. One possibility is that nouns are simply easier to learn (see Ellis & Beaton, 1993, but also Schwanenflugel, Stahl, & McFalls, 1997, for an alternative account). Another is that because nouns are more concrete than the other word types in in our dataset, they were easier to depict and define with in-text scaffolds, thus making scaffolding feature more effective for this word type.

Practical significance

The strength of the associations reported in this study was modest, that is odds ratios were lower than 1.68 (Chen et al., 2010). However, in terms of practical significance these associations may be meaningful. First, the literature on incidental vocabulary learning and reading development suggests that small effects, when cumulative, can have long term benefits over time (Anderson et al., 1988; Fukkink & de Glopper, 1998). Second, the associations reported here were captured in highly naturalistic settings and are potentially very conservative estimates, given the noise inherent in this type of dataset. In addition, vocabulary scaffolds embedded in digital supplements represent a relatively low-cost approach to support, as compared to more resource intensive approaches, such as tutoring or classroom-based instructions. In this context, our identification of positive associations suggests the potential utility of including scaffolds in digital supplements to support independent reading and, through that, both comprehension and vocabulary growth. Taken together, these observations lead us to recommend the inclusion of embedded scaffolds in digital reading materials to provide timely support.

Exploring the Contribution of Different Scaffold Components and their Combinations

Our exploratory analysis further suggests that a written definition alone, and especially in conjunction with pictorial support, is associated with a higher proportion of correct responses compared to other combinations of scaffold components, such as the combination of picture, pronunciation, written and spoken definitions. These findings are in line with, and extend, previous evidence indicating that children's narrative skills benefit more from the combination of verbal and non-verbal scaffolds, than from just verbal scaffolds alone (Pesco & Gagné, 2017), as well as evidence documenting the benefits of imagery training to children's reading comprehension (Francey & Cain, 2015; Joffe et al., 2007). Our descriptive analysis also suggests that the use of multiple scaffold components at a time is not necessarily associated with

higher performance. This can be interpreted in several ways. One possibility is that performance when using the scaffold was already high, leaving little room for improvement when additional spoken supports were used. Another possibility is that interacting with several components of the scaffold may increase cognitive load and distract students from the main task (Mayer & Moreno, 2003). This may explain why performance was less accurate when children used several scaffold components for the same item. An alternative account of these exploratory results is that children clicked on more components of the scaffold when they faced a particularly challenging item, suggesting that the benefits of scaffolds may vary by item difficulty. Future experimental work is needed to validate this pattern of results under more controlled conditions.

Exploring the Influence of Key Psycholinguistic Variables

Our exploratory analysis further investigated whether key psycholinguistic variables such as age of acquisition and concreteness modulated the association between use of scaffolds and reading comprehension. Our data suggest that they did not. Again, future experimental work is needed to test the reproducibility of this set of results under carefully controlled conditions. Information about any influence of the psycholinguistic properties of words in relation to both the use of scaffolds and their associations with reading performance is crucial to inform both practice and theory.

Limitations and Future Research

This study comes with limitations and suggestions for future research, additional to those already discussed. First, causal inference is beyond the scope of this analysis, due to the observational nature of the dataset and lack of experimental manipulation. It is however plausible that the nature of the associations described in this work are causal; that is, the use of the scaffold resulted in enhanced item accuracy. This possibility should be taken up by the research community and tested in future experimental work, in a cumulative science framework. Engagement and motivation may also play an important in reading performance and future work should take those factors into account when examining the role of scaffolding features. A strength of our approach was the interrogation of a substantial dataset of children's interactions with a digital reading environment providing the statistical power to examine several child characteristics. Concerns have been raised in relation to high-

powered studies. We note here that those studies are problematic in conjunction with null-hypothesis significance testing only when the sample is increased arbitrarily to achieve significance and/or p-values are the only criterion used to evaluate results. This does not apply to our study as we have taken the strength of the relationships into account and have exploited an existing big dataset. A large sample in this context provides more stable estimates, increasing the chance that these patterns of results will be observed in future studies (addressing current and critical concerns about the reproducibility of scientific findings). Of course, our findings are limited to the specific reading game under investigation; future research is needed to clarify whether these associations generalize to a different sample, context, type of text, and reading comprehension measure. Finally, the differential impact of the scaffolds across word types is an important area for future research to inform how to best scaffold the meaning of adjectives and adverbs, the impact of scaffolds for verbs and prepositions, and the role of other key psycholinguistic variables. This information is critical to develop appropriate scaffolds for a range of vocabulary.

Conclusions

Vocabulary scaffolding features may be promising tools to promote reading comprehension in general, and to reduce performance differences amongst diverse readers in digital environments. This work can be viewed as a proof of principle for the feasibility of collaborations between academics and industrial partners in the field of language and literacy research in the era of large app-generated data. Specifically, we have shown that such collaborations can be undertaken in accordance with academic standards and that they can contribute meaningfully to inform a broader audience of app developers, educators, and policy makers.

Acknowledgments

We sincerely thank all the participating schools, teachers, parents, and children for contributing data. We also thank the Literacy Acceleration through Scalable Research (LASeR) consortium and all the collaborators across the industrial and academic network. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Actions grant agreement No 857897.

CRediT Author Statement

L. Diprossimo: Conceptualization, Formal analysis, Data curation, Writing original draft, Visualization, Project administration; **A. Ushakova:** Conceptualization, Formal analysis, Writing original draft, Project administration; **J. Zoski:** Resources, Writing - review & editing; **H. Gamble:** Resources, Data curation, Writing - review & editing; **R. Irey:** Writing - review & editing; **K. Cain:** Conceptualization, Writing original draft, Supervision, Funding acquisition.

References

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*(3), 199–226. https://doi.org/10.1080/09588220802090246

Abu Rabia, S. (2019). The effect of degrees of bilingualism on metacognitive linguistic skills. *International Journal of Bilingualism*, *23*(5), 1064–1086. https://doi.org/10.1177/1367006918781060

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of english as a foreign language. *Instructional Science*, *35*(6), 499–517. https://doi.org/10.1007/s11251-007-9016-7

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. Wellcome open research, 4, 63. https://doi.org/10.12688/wellcomeopenres.15191.1

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading research quarterly*, 285-303.

Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children. *Educational Psychology Review*, *32*(3), 815–850. https://doi.org/10.1007/s10648-020-09527-z

Baker, L., & Cerro, L. C. (2000). *Assessing Metacognition in Children and Adults*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blything, L. P., Hardie, A., & Cain, K. (2020). Question asking during reading comprehension instruction: a corpus study of how question type influences the linguistic complexity of primary school students' responses. *Reading Research Quarterly*, *55*(3), 443–472. https://doi.org/10.1002/rrq.279

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Cain, K., & Oakhill, J. V. (2018). Vocabulary development and reading comprehension: a reciprocal relationship. *Oxford University Press.*, 1. http://oxford.ly/wordgap

Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, *30*(3), 681–694. https://doi.org/10.1017/S0305000903005713

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—simulation and Computation®*, *39*(4), 860-864.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: a meta-analysis. *Review of Educational Research*, *90*(4), 499–541. https://doi.org/10.3102/0034654320933544

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978. https://doi.org/10.1037/0012-1649.38.6.967

Clark, K. F., & Graves, M. F. (2005). Scaffolding students' comprehension of text. *The Reading Teacher*, *58*(6), 570–580. https://doi.org/10.1598/rt.58.6.6

Cohen, P. A., Kulik, J. A., & Kulik, C.L. C. (1982). Educational outcomes of tutoring: a meta-analysis of findings. *American Educational Research Journal*, *19*(2), 237–248. https://doi.org/10.3102/00028312019002237

Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science*, *30*(6), 433–464. https://doi.org/10.1023/A:1020516301541

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *Journal of Direct Instruction*, *1*(2), 137–149.

Dalton, B., Proctor, C. P., Uccelli, P., Mo, E., & Snow, C. E. (2011). Designing for

diversity: The role of reading strategies and interactive vocabulary in a digital reading environment for fifth-grade monolingual English and Bilingual students. *Journal of Literacy Research*, *43*(1), 68–100. https://doi.org/10.1177/1086296X10397872

De Sousa, I., & Oakhill, J. (1996). Do levels of interest have an effect on children's comprehension monitoring performance? *British Journal of Educational Psychology*, *66*(4), 471–482. https://doi.org/10.1111/j.2044-8279.1996.tb01213.x

Degener, S., & Berne, J. (2017). Complex questions promote complex thinking. *Reading Teacher*, *70*(5), 595–599. https://doi.org/10.1002/trtr.1535

Diprossimo, L., Cain, K., & Ushakova, A. (2021, June 9). Vocabulary Scaffolding Features and Young Readers' Comprehension of Digital Text: Insights from a Big Observational Dataset. https://doi.org/10.17605/OSF.IO/62C4Q

Duncan, L. G., Mcgeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology*, *107*(2), 209–238. https://doi.org/10.1111/bjop.12134

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. https://doi.org/10.1080/10888438.2013.774005

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

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559–617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Francey, G., & Cain, K. (2015). Effect of imagery training on children's comprehension of pronouns. *Journal of Educational Research*, *108*(1), 1–9. https://doi.org/10.1080/00220671.2013.824869

Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. Review of educational research, 68(4), 450-469.

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: a meta-analysis. In *Review of Educational Research: Vol. XX* (Issue X). https://doi.org/10.3102/0034654321998074

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural

partitioning. In S. A. Kuczaj (Ed.), *Language development: Language, thought, and culture (Vol. 2, pp. 301–334)*. Hillsdale, NJ: Erlbaum.

Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, *4*, 189–202.

Gonzalez, M. (2014). The effect of embedded text-to-speech and vocabulary ebook scaffolds on the comprehension of students with reading disabilities. *International Journal of Special Education*, *29*(3), 111–125.

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181–198. https://doi.org/10.1002/rrq.130

Hald, L. A., de Nooijer, J., van Gog, T., & Bekkering, H. (2016). Optimizing word learning via links to perceptual and motoric experience. *Educational Psychology Review*, *28*(3), 495–522. https://doi.org/10.1007/s10648-015-9334-2

Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*, *37*(2), 205–225. https://doi.org/10.1177/0142723716679956

Joffe, V. L., Cain, K., & Marić, N. (2007). Comprehension problems in children with specific language impairment: Does mental imagery training help? *International Journal of Language and Communication Disorders*, *42*(6), 648–664. https://doi.org/10.1080/13682820601084402

Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences*, *27*, 82–89. https://doi.org/10.1016/j.lindif.2013.07.002

Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Journal of Educational Technology & Society*, *11*(3), 114–126. https://www.jstor.org/stable/10.2307/jeductechsoci.11.3.114

Kim, N. J., Belland, B. R., & Walker, A. E. (2018). Effectiveness of computer-based scaffolding in the context of problem-based learning for stem education: bayesian meta-analysis. *Educational Psychology Review*, *30*(2), 397–429. https://doi.org/10.1007/s10648-017-9419-1

Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., & Gruelich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing*, *27*(2), 237–253. https://doi.org/10.1007/s11145-013-9440-9

Kirby, J. R., & Moore, P. J. (1987). Metacognitive awareness about reading and its relation to reading ability. *Journal of Psychoeducational Assessment*, *2*, 119–137.

Korat, O., & Shamir, A. (2012). Direct and indirect teaching: Using e-books for supporting vocabulary, word reading, and story comprehension for young children. *Journal of Educational Computing Research*, *46*(2), 135–152. https://doi.org/10.2190/EC.46.2.b

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lämsä, J., Hämäläinen, R., Aro, M., Koskimaa, R., & Äyrämö, S. M. (2018). Games for enhancing basic reading and maths skills: A systematic review of educational game design in supporting learning by people with learning disabilities. *British Journal of Educational Technology*, *49*(4), 596–607. https://doi.org/10.1111/bjet.12639

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, *50*(2), 151–169. https://doi.org/10.1002/rrq.99

Language and Reading Research Consortium, Currie, N. K., & Muijselaar, M. M. L. (2019). Inference making in young children: The concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology*, *111*(8), 1416–1431. https://doi.org/10.1037/edu0000342

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. In *American Educational Research Journal* (Vol. 47, Issue 3). https://doi.org/10.3102/0002831209355469

Liu, Y. T., & Leveridge, A. N. (2017). Enhancing L2 vocabulary acquisition through implicit reading support cues in e-books. *British Journal of Educational Technology*, *48*(1), 43–56. https://doi.org/10.1111/bjet.12329

Lüdecke, D. (2018). ggeffects: tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M. A., Guerrero-Puerta, L., Aguilar-Parra, J. M., Trigueros, R., & Alias, A. (2021). Between level up and game over: A systematic literature review of gamification in education. *Sustainability*, *13*(4), 1–14. https://doi.org/10.3390/su13042247

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, *33*(5), 403–423. https://doi.org/10.1111/jcal.12197

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/s0959-4752(02)00016-6

Mayer, R. E. (2020). Where is the learning in mobile technologies for learning?. Contemporary Educational Psychology, 60, 101824. https://doi.org/10.1016/j.cedpsych.2019.101824

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *American Psychological Association, Inc*, *86*(3), 389–401. http://libproxy.library.unt.edu:2067/ehost/pdfviewer/pdfviewer?vid=1&sid=cace

cf36-cd5b-4ffe-9996-9b0332df5fb7%40pdc-v-sessmgr05

Mckeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension: a comparison of instruction for strategies and content approaches. *Reading Research Quaterly*, *44*(3), 218–253. https://doi.org/10.1598/RRQ.44.3.1

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, 112, 104092. https://doi.org/10.1016/j.jml.2020.104092

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Nassaji, H. (2003). L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in l2 lexical inferencing. *Tesol Quarterly*, *37*(4), 645–670. https://doi.org/10.2307/3588216

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. https://doi.org/10.1080/10888438.2010.529219

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Pesco, D., & Gagné, A. (2017). Scaffolding narrative skills: a meta-analysis of instruction in early childhood settings. *Early Education and Development*, *28*(7), 773–793. https://doi.org/10.1080/10409289.2015.1060800

Proctor, C. P., Dalton, B., & Grisham, D. L. (2007). Scaffolding English language learners and struggling readers in a universal literacy environment with embedded strategy instruction and vocabulary support. *Journal of Literacy Research*, *39*(1), 71–93. https://doi.org/10.1080/10862960709336758

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: a latent change score modeling study. *Child Development*, *86*(1), 159–175. https://doi.org/10.1111/cdev.12292

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures

as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, *38*(2), 109–117. https://doi.org/10.1016/j.cedpsych.2012.12.001

Sadeghi, K., & Khezrlou, S. (2012). Glossing mode in self-regulated vocabulary learning, and its relationship with gender, age, and field of study. *Journal of Asia TEFL*, *9*(3), 51–74.

Salmerón, L., & García, V. (2011). Reading skills and children's navigation strategies in hypertext. *Computers in Human Behavior*, *27*(3), 1143–1151. https://doi.org/10.1016/j.chb.2010.12.008

Sampayo-Vargas, S., Cope, C. J., He, Z., & Byrne, G. J. (2013). The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Computers & Education*, *69*, 452-462.

Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, *29*(4), 531–553. https://doi.org/10.1080/10862969709547973

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning*, *59*(3), 623–646. https://doi.org/10.1111/j.1467-9922.2009.00519.x

Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, *27*(1), 78–102. https://doi.org/10.1093/applin/ami046

Wolf, M., & Katzir-Cohen, T. (2009). Reading fluency and its intervention. *Scientific Studies of Reading*, *5*(3), 211–239. https://doi.org/10.1207/S1532799XSSR0503

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wu, J. Y. (2014). Gender differences in online reading engagement, metacognitive strategies, navigation skills and reading literacy. *Journal of Computer Assisted Learning*, *30*(3), 252–271. https://doi.org/10.1111/jcal.12054

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, *24*(1), 39–58. https://doi.org/10.1080/09588221.2010.523285

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning*, *34*(5–6), 751–777. https://doi.org/10.1080/09588221.2019.1640745

Zucker, T. A., Justice, L. M., Piasta, S. B., & Kaderavek, J. N. (2010). Preschool teachers' literal and inferential questions and children's responses during whole-class shared reading. *Early Childhood Research Quarterly*, *25*(1), 65–83. https://doi.org/10.1016/j.ecresq.2009.07.001

Chapter 5: General Discussion

The overarching aim of this thesis was to examine the factors that influence word learning from written contexts in early childhood, particularly during contemporary reading experiences which include, not only print books, but also digital books and digital reading instructional supports. The first objective was the investigation of any unique or mediated effects of book format (print vs digital) on 4- to 5-year-olds' word learning during shared reading with a caregiver (Chapter 2). The second objective was to assess the effects of the book format on caregivers' verbal and non-verbal scaffolding during shared reading and to test the generalisability of the fine-tuning hypothesis to the shared reading context (Chapter 3). The third objective was to explore the associations between the use of vocabulary scaffolds and reading comprehension for 5- to 8-year-olds reading independently on screen (Chapter 4). Throughout the studies, particular attention was given to elucidating the influence of individual child characteristics, including vocabulary knowledge, reading skills, executive functions, bilingualism, gender, and age. In the next sections, I will discuss and integrate key study findings with the extant literature and theoretical models of word learning and reading. The broader implications of this work for theory and practice are also elaborated. I will close by noting the limitations of the current research and highlighting opportunities for future research in this area.

Effects of Book Format

The studies reported in Chapters 2 and 3 provide complementary insights into the effects of book format during shared reading, by examining its effects on both children's word learning (Chapter 2) and caregivers' provision of scaffolds (Chapter 3). For children's word learning, a complex pattern of findings emerged, which may, at least in part, explain the contradictory findings synthesised in prior meta-analytic research (Furenes et al., 2021). Book format affected only certain aspects of word learning and this effect was contingent on child characteristics. Specifically, digital-based shared reading was positively associated with semantic recognition in children with higher executive function, and negatively associated with semantic recognition in children with lower executive function. For caregivers' scaffolds, the results were more straightforward. The digital format had a negative impact on caregivers' verbal and

gestural scaffolds, extending prior findings reporting lower communicative initiations, responses, and expanding talk by caregivers when reading with their child using digital compared to print format (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020). When integrating the findings reported in Chapter 2 and 3, it is noteworthy that the differences in caregivers' scaffolds related to book-format were not directly reflected in children's word learning. Indeed, there was evidence that the effect of book format on word learning was conditional on executive function and that the effect of verbal scaffolds on word learning was conditional on children's prior knowledge of target words (see next section for further discussion on fine-tuning). Taken together, the findings reported in Chapters 2 and 3 provide support for the *dynamic model of children's e-reading* (Kucirkova, 2019) which emphasises the importance of considering child characteristics and caregivers' scaffolding when studying the effect of book format on word learning in the context of shared reading. They also align with the capacity model (Fisch, 2000) and the differential susceptibility to media effects model (Valkenburg & Peter, 2013), which underscore the importance of considering child characteristics and social-context variables when studying learning from different media.

The cognitive theory of multimedia learning also provides a useful framework to interpret the word-learning results; within this framework, executive functions can be understood to have acted as a bottleneck for learning from the digital format, potentially due to the cognitive load associated with handling and operating the digital device (Courage, 2019; Honma et al., 2022; Mayer, 2003; Mayer & Moreno, 2003). For children with higher executive function, the digital format had a positive impact on semantic recognition. This advantage might be attributable to the novelty of the digital device to our participants, whose caregivers reported substantially less experience with digital-based, compared to print-based, shared reading (Flewitt et al., 2015; O'Toole & Kannass, 2018; Reich et al., 2019). Novelty may also explain differences in caregiver scaffolding behaviours across book formats, as one could speculate that being less familiar with the digital reading format could interfere with the flow of interactions, which are more established in the more familiar print format. Another important aspect to consider in interpreting the effect of book format on caregivers' scaffolds is their perceptions of child engagement across book formats, and also

caregivers' preference for print over digital books, which may have influenced their scaffolding behaviour (Kucirkova, 2019).

Effects of Scaffolds

Caregivers' modulation of scaffolds and its effect on children's word learning during shared reading were examined in Chapter 3, while the uptake and effectiveness of vocabulary scaffolding features in a digital reading supplement were investigated in Chapter 4. Caregivers modulated verbal, but not gestural, scaffolds as a function of children's individual lexical knowledge, and this modulation supported children's semantic recall. These results replicate and extend prior research (Donnellan et al., 2023; Leung et al., 2021; Shi et al., 2022) and provide support for the generalisation of the fine-tuning hypothesis (Leung et al., 2021). From a theoretical perspective, these findings demonstrate caregivers' sophisticated ability to adjust their communication to their children's individual lexical knowledge and its effectiveness in supporting children's word learning, consistently with the notions of scaffolding (Wood et al., 1976) and zone of proximal development (Vygotsky, 1978). Considering the lexical knowledge of their child, caregivers offer tailored support that facilitates the acquisition of the semantic features of new words during shared reading.

The findings for digital scaffolds (Chapter 4) suggest that their use was positively associated with reading comprehension in a digital reading supplement, supporting the utility of embedding multimodal supports in digital reading materials. These results extend previous meta-analytic evidence on university students reading a foreign language (Abraham, 2008). However, our results contrast with meta-analytic evidence showing a null or negative effect of digital scaffolds on children's story comprehension (Furenes et al., 2021). These differences between studies may be due to the longer texts used in the storybooks examined in other research, and possibly from the co-reading context that was the focus of the meta-analysis. Theoretically, the positive association found between the use of multimodal vocabulary scaffolds and reading comprehension is in line with both the dual coding theory and the cognitive theory of multimedia learning, which both predict learning advantages associated with pictorial supports.

It is also worth noting that, in our study, scaffolds were available to learners on an as-needed basis. We found that the use of scaffolds was on average lower than

desired, suggesting the need to identify strategies to increase the use of scaffolds. Both the uptake and effectiveness of scaffolds were modulated by child characteristics. These findings are discussed in greater detail in the next section on the influence of child characteristics.

Taken together, the findings reported in Chapters 3 and 4 advance our understanding of scaffolds provided by caregivers and educational technology. They provide insights into the features that make scaffolds most effective. For example, caregivers' scaffolds were most effective when tailored to the child's individual word knowledge. Digital scaffolds were most effective for lower-achieving groups. They also suggest that the use of digital scaffolds is limited during independent reading, and point to the importance of adult mediation as well as the provision of digital scaffolds in an adaptive manner, as inspired by caregivers fine-tuning behaviours.

Effects of Child Characteristics

Across empirical studies, a range of child characteristics were considered, including prior vocabulary knowledge, reading skills, executive functions, bilingualism, gender, and age. I summarize and discuss the results relating to these next.

In the context of shared reading, there was evidence that prior vocabulary knowledge supported word learning, which is in line with prior research (Sénéchal et al., 1995). One explanation for this effect is that prior knowledge provides a foundation for learning, perhaps by freeing up cognitive resources. It is also possible that greater knowledge of semantic categories eases the acquisition of new words (Borovsky et al., 2016). There was also evidence that executive function uniquely predicted semantic recall from shared reading, which is in line with accumulating evidence (Hadley et al., 2021; Kapa & Erikson, 2020). This finding provides support for models of discourse comprehension that acknowledge the contribution of higher-level cognitive skills and processing resources (Cain, Oakhill, et al., 2004; Y. S. Kim, 2020). With respect to the effects of gender, boys significantly outperformed girls in word learning from shared reading, which is at odds with prior studies suggesting the opposite pattern of results, with girls being more accurate on a range of language and literacy measures (Eriksson et al., 2012; Frank et al., 2017; S. Logan & Johnston, 2010). Yet some recent findings are in line with ours and suggest it is important to consider the specific contexts where gender differences are observed (Bergman Deitcher et al., 2019; S. Logan & Johnston,

2010). Importantly, the gender differences in word learning reported in Chapter 2 were unlikely to be due to difference in caregivers' scaffolding. Indeed, the preliminary analyses reported in Chapter 3 indicated that caregivers' scaffolds were not conditional on child gender.

In the context of digital reading instruction support, child characteristics were also important and predicted the use of scaffolds. Children with lower literacy skills, beginner readers, girls, and students who reported as bilingual were more likely to use the scaffold. The results for struggling and beginner readers were particularly encouraging as they suggested that those learners have sufficient meta-cognitive skills to seek support when using the digital reading supplement. The results for girls and bilingual students are in line with research documenting enhanced metacognitive skills in girls (Wu, 2014) and bilingual readers (Abu Rabia, 2019). Further, the association between the use of scaffolds and reading comprehension was stronger for children with lower literacy skills, beginner readers, boys and bilingual students. This aligns with and extends the findings of a meta-analysis on the impact of similar scaffolds when acquiring novel words in a foreign language (Yun, 2011), showing that the impact of scaffolds was greater for beginner, compared to advanced, learners.

An integrated reading of the gender-related difference emerging from the three empirical chapters supports the notion that gender differences are context sensitive. In younger children and in the context of shared reading, boys performed better at word learning, particularly so in the semantic recall task. This task likely requires greater confidence, which is typically reported to be higher in boys compared to girls (Clark, 2011). In this context, boys' greater confidence might explain their better performance. In older children and in the context of independent reading on screen, boys made less use of optional vocabulary scaffolds and were less accurate than girls in the reading comprehension task. Together, these findings suggest that boys' greater confidence can aid or hinder their performance in different situations. Confidence should be measured directly in future studies to understand the potential mechanisms driving gender-difference in word learning in a range of educational contexts.

Methodological innovations

The thesis as a whole covers a range of interconnected research questions to clarify, replicate and extend prior findings, as well as to explore new hypotheses. Each study included methodological innovations. In Chapter 2, I addressed the contradictory findings on the effect of book format on children's word learning in the context of shared reading synthesised in the meta-analyses by Furenes et al. (2021). In a pre-registered study, I introduced several methodological improvements that enabled to quantify any unique or mediated effect of book format on word learning from shared reading in 4- to 5-year-olds. The methodological improvements from previous work included a within-subjects design and large sample of dyads, which increased the statistical power relative to previous research; having the same reader to narrate the text across book formats; collecting multiple outcome measures to assess both phonological and semantic features of word learning; and a semi-naturalistic paradigm to support generalisation outside the laboratory.

In Chapter 3, I extend prior research on the effect of book format on the quality of interactions during shared reading (Korat & Or, 2010; Munzer et al., 2019; Ozturk & Hill, 2020). Prior findings have evidenced differences in caregivers' communicative initiations, responses and expanding talk (Korat & Or, 2010), and social reciprocity between the parent-toddler dyads (Munzer et al., 2019) as a function of book format. I pre-registered and conducted further analyses on the dataset and corpus of video recordings collected in the context of the shared reading study reported in Chapter 2. This enabled an extension of prior research by considering the effect of book format on a range of verbal and non-verbal scaffolds provided by caregivers during shared reading. In the study reported in Chapter 3, I also tested the generalisability of the fine-tuning hypothesis (Leung et al., 2021) to the shared reading context. Recent research has suggested that caregivers adapt their speech and language to their child's lexical knowledge during toy play and a tablet-based reference game and that this supports children's word learning (Donnellan et al., 2023; Leung et al., 2021), but it remained unclear whether this mechanism would generalise to the shared reading context and extend to an older age range. The current work provides empirical evidence for the robustness and generalisability of fine-tuning across socio-interactional contexts and an extended age range.

In Chapter 4, the association between the use of vocabulary scaffolding features embedded in a digital reading environment and 5- to 8-year-olds' reading comprehension was examined. Previous work has shown that vocabulary scaffolds support word learning, but can interfere with the comprehension of the text (Abraham, 2008; Furenes et al., 2021). By leveraging a big naturalistic dataset of user log files, this was one of the first studies that examined several child and word characteristics that could potentially moderate the effectiveness of scaffolds.

Practical implications

These findings still require replication; if found, there would be several practical implications. First, the use of print-based reading material should be favoured in children with still immature executive functions. Second, tailoring support to each child's individual word knowledge is a promising avenue to support word learning: Fine-tuning can usefully inform the design of individualised scaffolding in digital media. Finally, vocabulary scaffolds should be embedded in digital reading material to provide timely support to diverse learners and promote both reading comprehension and vocabulary growth.

Limitations and Future Directions

This work comes with limitations and suggestions for future research, in addition to those already discussed. One limitation is that not all the studies reported in this thesis considered different word types. Whilst the study reported in Chapter 4 considered different word types, including nouns, verbs, adjectives, adverbs, and prepositions, the studies reported in Chapters 2 and 3 included only concrete nouns of animals and tools. It is plausible that certain word types benefit more from specific types of scaffold than others, as illustrated by the study reported in Chapter 4 and prior research. For instance, gestures are more beneficial for learning verbs referring to actions than other types of words (De Nooijer et al., 2013; de Nooijer et al., 2014). Abstract words play a critical role in the development of academic vocabulary, yet their acquisition is still poorly understood - presumably because measuring abstract word knowledge is challenging in early development (Hadley et al., 2016; Hadley & Dickinson, 2020; Sadoski & Lawrence, 2023). Future research should consider a range of word types, including adjectives, verbs, adverbs, and prepositions, as well as abstract vocabulary, to clarify the mechanisms that underpin the acquisition of words

other than concrete nouns. This will significantly advance our understanding of vocabulary acquisition, as any theoretical model of word learning needs to accommodate different word types (Tomasello, 2003).

Another limitation that should be noted, is the relatively limited exposure to target words in the studies reported in Chapters 2 and 3, where a single shared reading session was considered. Therefore, the acquisition of expressive vocabulary knowledge was limited. Learning the phonological form and particularly the meaning of new words requires multiple exposures over time and across different contexts, as predicted by the *word experience model* of adult readers (Reichle & Perfetti, 2003; Verhoeven & Perfetti, 2011). Another limitation is represented by the short delay at which learning was assessed. As such, future studies should include exposure to words across multiple sessions and contexts and also word retention assessments, to better understand the mechanisms that support word learning over longer, and perhaps more realistic, time scales. In this context, a consideration of the effect of massed and distributed practice, as well as retrieval practice, may provide useful insights (Childers & Tomasello, 2002; Laurence et al., 2020).

Gender differences were evident in these studies. While greater use of digital scaffolds in girls compared to boys (Chapter 4) is in line with the literature on gender differences in meta-cognition and help-seeking behaviour (Thompson et al., 2012; Wu, 2014), the word learning advantage for boys (Chapter 2) is at odd which much prior literature (Eriksson et al., 2012; Frank et al., 2017; S. Logan & Johnston, 2010), supporting the view that, at least in some cases, gender differences are context-dependent. This observation underscores the importance of investigating gender differences in the uptake of scaffolds and learning from different media and contexts.

These findings are also limited to the cultural context of our participating sample. Further investigation is needed to generalise these findings to different cultural contexts. This is particularly important as these results may vary as a function of experience with digital media (Reich et al., 2019), which has been reported to differ even between European (geographically intended) countries (Gracia et al., 2020).

Conclusions

In an increasingly digitised society, children's literacy experiences are rapidly changing. In this context, the need to build the evidence base to inform early literacy

policy and practice is widely recognised. This thesis has taken different angles to shed further clarity on how different media and media affordances influence word learning from written context in early childhood. Three key conclusions can be drawn from the series of studies presented here.

The first is the critical role of child characteristics. Child characteristics were shown to modulate the effects of both book format and scaffolding features. It follows that child characteristics need to be considered in future research and when developing recommendations for stakeholders. Here it is important to acknowledge the complexities associated with conveying nuanced messages to the general public and stress the importance of interdisciplinary collaborations. Experts in communication should assist in the effort to make research evidence accessible to a broader audience.

The second key message that emerged, is that an in-depth understanding of caregivers' scaffolding and its relation to children's learning has the potential to inform the design of digital technology. For instance, caregivers' fine-tuning can guide the design and implementation of individualised digital scaffolds. Individualisation of content is a powerful affordance of digital tools. This affordance is promising to support each child in reaching their full developmental potential, but also impose careful reflection on how it can be sensibly implemented to narrow, rather than exacerbate, differences between children.

The third and concluding remark is a plea for sustained partnerships between developmental scientists and app developers. In the final study, we demonstrated how industry and researchers can work together to advance knowledge with due independence and mitigating conflicts of interest. Our findings can now be fed into enhancements of digital scaffolds. Responsible collaborative partnerships, such as these, will help to make the most out of the many opportunities that digital tools bring for understanding and supporting child development.

Supporting Information

Supplementary Materials for Chapter 2

Table S 1 *Sociodemographic Characteristics of Participants*

	Children	Caregivers
Age	Range 48 to 71 months M = 57.68 SD = 7.24	Range 29 to 47 years M = 37.76 SD = 3.93
Sex	Girls 57% Boys 43%	Women 83% Men 6% Did not state 11%
Education	Day Care Facility 1% Nursery 14% Pre-school 18% School 53% Did not state 14%	Secondary School 3% High School Diploma 12% Undergraduate/Master Degree 69% PhD 6% Did not state 10%
Socio Economic Status	Decile 1-5: 29% Decile 6-10: 57% Post code not stated 14%	
Child experience with print-based shared reading	Daily 86% Several Times a Week 2% Once or Twice a Week 1% More Than Once a Month 0% Never 0% Did not state 11%	
Child experience with print books alone	Daily 38% Several Times a Week 28% Once or Twice a Week 11% More Than Once a Month 6% Never 6% Did not state 11%	
Child experience with digital-based shared reading	Daily 1% Several Times a Week 3% Once or Twice a Week 8% More Than Once a Month 9% Never 68% Did not state 11%	

Child experience with digital books alone	Daily 2% Several Times a Week 0% Once or Twice a Week 7% More Than Once a Month 10% Never 70% Did not state 11%

Note. Socio economic condition was obtained by linking the postcode of participants to the index of multiple deprivation expressed in decile. Child experience with different book format in joint and solo settings was collected through caregiver report.

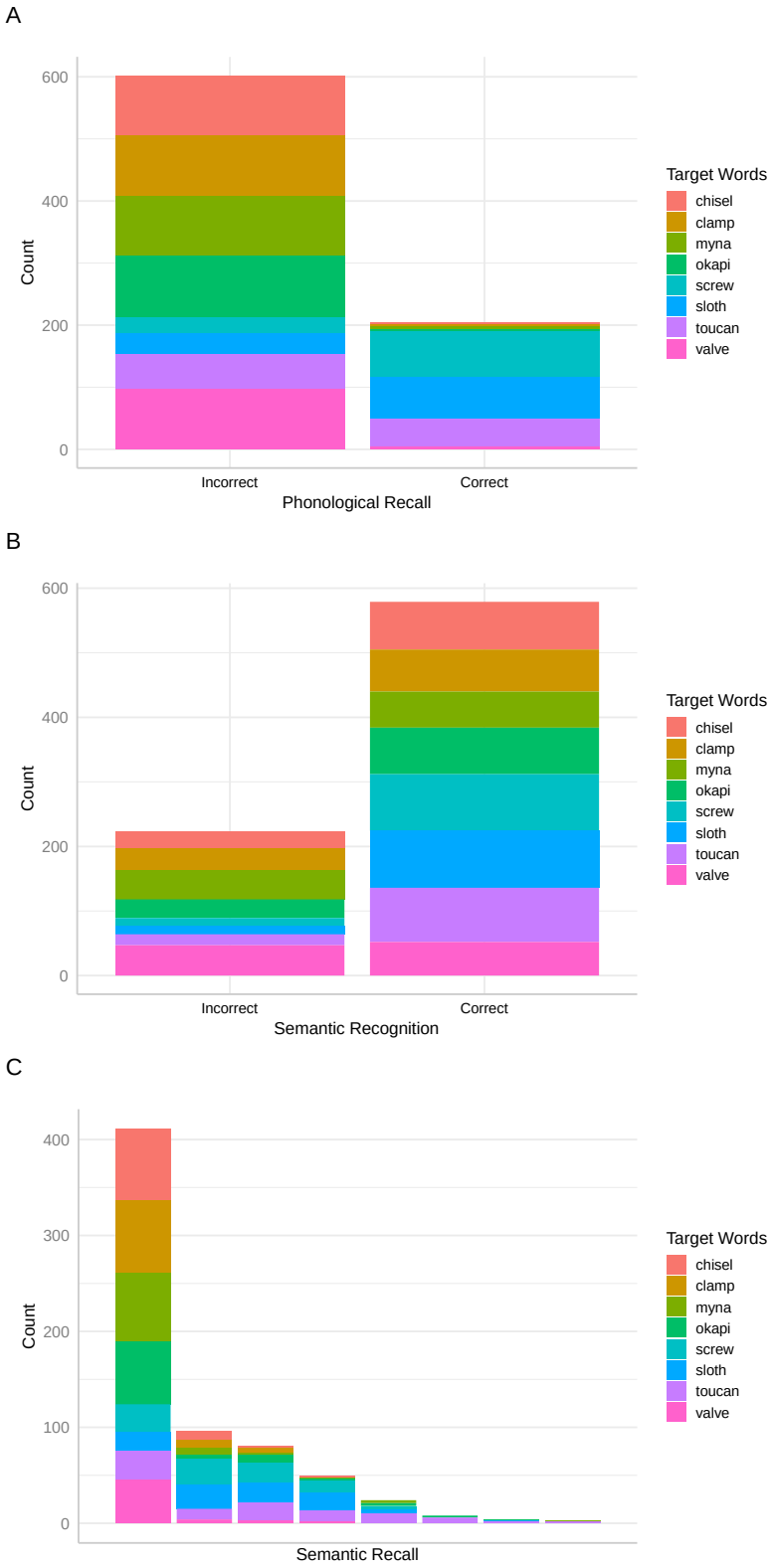
Table S 2 *Means, standard deviations, and correlations with confidence intervals, prior to conducting Principal Component Analyses on Executive Functions*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. age (months)	57.77	7.24								
2. vocabulary (scaled score)	10.15	2.81	.03 [-.04, .10]							
3. visual attention (scaled score)	11.23	2.96	.23** [.16, .29]	.12** [.05, .19]						
4. forward digit span	11.06	2.72	.49** [.44, .54]	.06 [-.01, .13]	.40** [.34, .45]					
5. backward digit span	5.68	2.62	.46** [.41, .52]	.29** [.23, .35]	.33** [.27, .39]	.44** [.38, .50]				
6. global executive function (t-score)	51.01	11.23	-.03	-.10**	-.13**	-.05	-.34**			

			[-.10, .04]	[-.17, -.03]	[-.20, -.06]	[-.13, .02]	[-.40, -.28]			
7. phonological recall	0.25	0.44	.07* [.00, .14]	.10** [.03, .16]	.05 [-.02, .12]	.07* [.00, .14]	.07* [.00, .14]	-.03 [-.10, .04]		
8. semantic recall	0.87	1.35	.09* [.02, .17]	.19** [.12, .27]	.01 [-.06, .09]	.09* [.02, .17]	.15** [.07, .22]	-.05 [-.13, .03]	.46** [.40, .51]	
9. semantic recognition	0.72	0.45	.01 [-.06, .08]	.15** [.08, .22]	.04 [-.03, .11]	.07* [.00, .14]	.09* [.02, .15]	-.03 [-.10, .04]	.28** [.22, .34]	.23** [.16, .30]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

Figure S 1 *Accuracy Across Word Learning Measures by Target Word*



Note – A) Phonological Recall; B) Semantic Recognition; C) Semantic Recall.

Table S 3 *Means (and Standard Deviations) of Word Learning Outcomes by Sex*

Sex	***Phonological recall***	***Semantic Recall***	***Semantic Recognition***
Girls	0.23 (0.42)	0.69 (1.15)	0.69 (0.46)
Boys	0.28 (0.45)	1.10 (1.55)	0.77 (0.42)

Table S 4 *Results of the Models Estimating the Effect of Book Format, Child Characteristics, and Their Interaction, on Phonological Recall*

	Phonological Recall											
	M_0				M_1				M_2			
Predictors	*OR*	*SE*	*CI*	*p*	*OR*	*SE*	*CI*	*p*	*OR*	*SE*	*CI*	*p*
(Intercept)	0.14	0.05	0.07 – 0.27	**<0.001**	0.13	0.05	0.07 – 0.27	**<0.001**	0.12	0.06	0.04 – 0.30	**<0.001**
Vocabulary	1.08	0.03	1.01 – 1.14	**0.017**	1.08	0.03	1.01 – 1.14	**0.017**	1.09	0.05	1.00 – 1.19	**0.045**
Executive Function	1.15	0.11	0.95 – 1.40	0.147	1.15	0.11	0.95 – 1.40	0.147	1.16	0.15	0.90 – 1.49	0.265
Age (months)	1.10	0.10	0.92 – 1.32	0.297	1.10	0.10	0.92 – 1.32	0.297	1.10	0.10	0.92 – 1.32	0.298
Sex [M]	1.40	0.23	1.01 – 1.94	**0.045**	1.40	0.23	1.01 – 1.94	**0.045**	1.40	0.23	1.01 – 1.94	**0.045**
Book format [digital]					1.01	0.17	0.73 – 1.39	0.949	1.36	0.91	0.37 – 5.02	0.642
Book format [digital] * Vocabulary									0.97	0.06	0.86 – 1.10	0.645
Book format [digital] * Executive Function									0.99	0.17	0.71 – 1.39	0.975
Random Effects												
σ^2	3.29				3.29				3.29			
τ_{00}	0.00 $_{child\_ID}$				0.00 $_{child\_ID}$				0.00 $_{child\_ID}$			
τ_{11}	0.00 $_{child\_ID.book\_format}$				0.00 $_{child\_ID.book\_format}$				0.00 $_{child\_ID.book\_format}$			
ρ_{01}												
N	100 $_{child\_ID}$				100 $_{child\_ID}$				100 $_{child\_ID}$			

Observations	807	807	807
Marginal R^2 / Conditional R^2	0.034 / NA		

Table S 5 *Results of the Models Estimating the Effect of Book Format, Child Characteristics, and Their Interaction, on Semantic Recognition*

	Semantic recognition											
	M_0				**M_1**				**M_2**			
Predictors	*OR*	*SE*	*CI*	*p*	*OR*	*SE*	*CI*	*p*	*OR*	*SE*	*CI*	*p*
(Intercept)	0.62	0.23	0.29 – 1.29	0.200	0.56	0.22	0.26 – 1.20	0.138	0.53	0.25	0.21 – 1.33	0.178
Vocabulary	1.13	0.04	1.06 – 1.21	**<0.001**	1.13	0.04	1.06 – 1.21	**<0.001**	1.14	0.05	1.04 – 1.24	**0.004**
Executive Function	1.26	0.15	1.00 – 1.59	0.053	1.25	0.15	0.99 – 1.58	0.063	1.07	0.15	0.82 – 1.41	0.609
Age (months)	0.90	0.10	0.72 – 1.13	0.365	0.90	0.10	0.73 – 1.13	0.372	0.91	0.10	0.73 – 1.13	0.372
Sex [M]	1.92	0.40	1.28 – 2.88	**0.002**	1.93	0.40	1.29 – 2.88	**0.001**	1.94	0.40	1.29 – 2.89	**0.001**
Book format [digital]					1.19	0.22	0.82 – 1.71	0.362	1.31	0.83	0.38 – 4.55	0.672
Book format [digital] * Vocabulary									0.99	0.06	0.88 – 1.12	0.920
Book format [digital] * Executive Function									1.41	0.25	1.00 – 1.99	**0.047**
Random Effects												
σ^2	3.29				3.29				3.29			

τ_{00}	0.22 $_{child\_ID}$	0.22 $_{child\_ID}$	0.22 $_{child\_ID}$
τ_{11}	0.28 $_{child\_ID.book\_format}$	0.26 $_{child\_ID.book\_format}$	0.16 $_{child\_ID.book\_format}$
ρ_{01}	0.25 $_{child\_ID}$	0.51 $_{child\_ID}$	0.75 $_{child\_ID}$
ICC	0.06	0.06	0.06
N	100 $_{child\_ID}$	100 $_{child\_ID}$	100 $_{child\_ID}$
Observations	802	802	802
Marginal R^2 / Conditional R^2	0.068 / 0.126	0.070 / 0.128	0.080 / 0.138

Table S 6 *Data-Driven Model M_3 With Non-Significant Interaction Term Pruned*

		Semantic Recognition		
Predictors	*Estimate*	*SE*	*CI*	*p*
(Intercept)	-0.60	0.39	-1.36 – 0.15	0.117
Vocabulary	0.13	0.04	0.06 – 0.20	**<0.001**
Book format [digital]	0.21	0.18	-0.15 – 0.57	0.259
Executive Function	0.07	0.14	-0.20 – 0.34	0.595
Age (months)	-0.10	0.11	-0.32 – 0.12	0.372
sex [M]	0.66	0.21	0.26 – 1.06	**0.001**
Book format [digital] * Executive Function	0.34	0.17	0.01 – 0.68	**0.045**
Random Effects				
σ^2	3.29			
$\tau_{00\ child\_ID}$	0.22			
$\tau_{11\ child\_ID.book\_format.M}$	0.16			
$\rho_{01\ child\_ID}$	0.76			
ICC	0.06			
$N_{\ child\_ID}$	100			
Observations	802			
Marginal R^2 / Conditional R^2	0.080 / 0.139			

Supplementary Materials for Chapter 3

Figure S 2 *Pattern of Missingness in the Dataset by Variables*

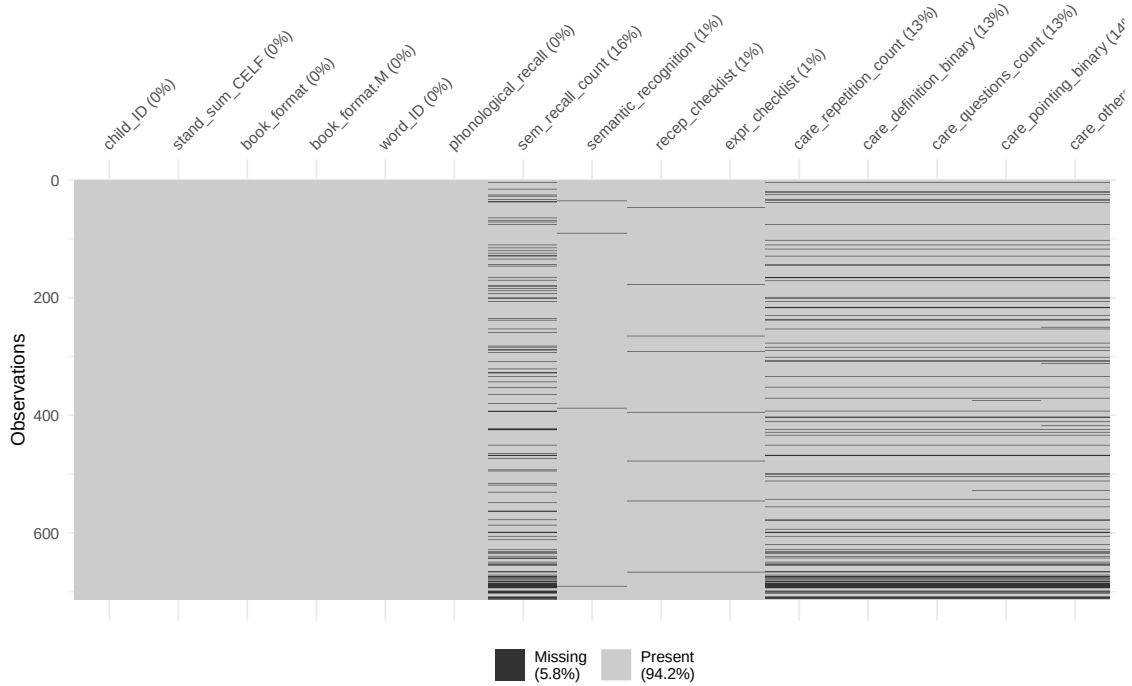


Figure S 3 *Pattern of Missingness by Variables and Child IDs*

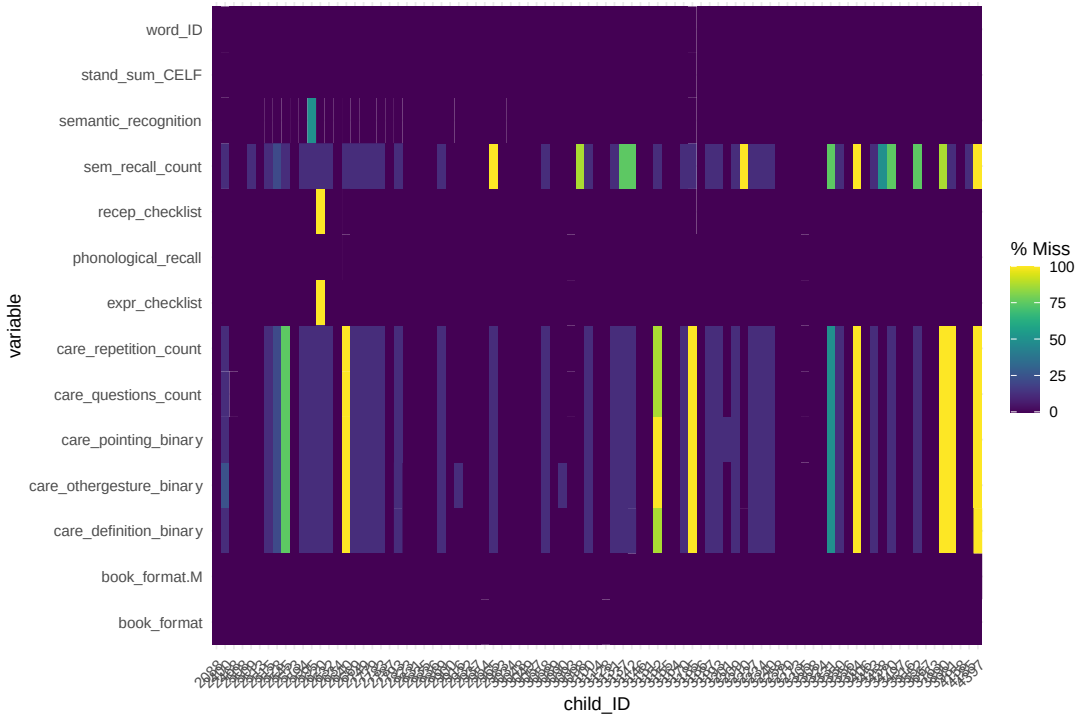


Table S 7 *Frequency and Age of Acquisition of Target Words and Accompanying Adjectives*

	Story A			Story B	
	Frequency	AoA		Frequency	AoA
Target words					
okapi	2.23	11.22	clamp	2.71	10.89
sloth	3.38	8.37	valve	3.31	10.78
myna	2.23	9.5	chisel	2.93	10.53
toucan	3.71	8.69	screw	3.82	6.65
Adjectives					
striped	3.08	4.65	wooden	4.34	5.89
furry	4.62	5.72	shiny	4.32	5.05
dark	5.28	3.74	sharp	4.82	6.11
colourful	4.82	4.89	pointy	4.63	7.39

Note. Frequency of SUBLEX. Values lower than 3.5 are considered infrequent in the corpus. AoA = Age of Acquisition

Table S 8 *Codebook for Behavioural Coding of Caregivers' Scaffolds during Shared Reading Interactions*

Scaffold code	Description
Repetition	How many times each target word is produced in extra textual talk by the caregiver.
Definition	Whether definitional information is given in extratextual talk by the caregiver, including synonyms and perceptual or conceptual information.
Comment	How many times the caregiver makes a comment which is related to the target word. E.g., look at that!
Question	How many times the caregiver asks a question related to the target word. E.g., can you find the [x]?
Pointing	Whether the caregiver points to the target word.
Iconic gesture	Whether the caregiver gestures to illustrate the target word meaning.

Table S 9 *Means, standard deviations, and correlations with confidence intervals*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. stand_sum_CELF	10.37	2.68										
2. recep_checklist	0.26	0.44	-.04 [-.12, .04]									
3. expr_checklist	0.39	0.49	-.00 [-.08, .08]	.10* [.01, .18]								
4. care_repetition_count	1.21	1.35	.08 [-.00, .16]	-.03 [-.11, .06]	-.10* [-.18, -.01]							
5. care_definition_binary	0.42	0.49	.01 [-.07, .09]	-.12** [-.20, -.03]	-.18** [-.26, -.10]	.32** [.25, .40]						

6. care_questions_count	1.10	1.47	.13**	-.00	-.07	.49**	.27**				
			[.04, .21]	[-.09, .08]	[-.16, .01]	[.42, .55]	[.19, .35]				
7. care_pointing_binary	0.71	0.45	.03	-.04	-.00	.19**	.14**	.19**			
			[-.06, .11]	[-.12, .05]	[-.09, .08]	[.11, .27]	[.06, .22]	[.11, .27]			
8. care_other gesture_binary	0.11	0.31	-.01	.04	-.07	.21**	.37**	.16**	.12**		
			[-.09, .08]	[-.04, .12]	[-.15, .02]	[.13, .29]	[.30, .44]	[.08, .24]	[.04, .20]		
9. phonological_recall	0.27	0.44	.09*	.16**	.55**	-.05	-.27**	-.00	-.03	-.11**	
			[.01, .17]	[.08, .24]	[.49, .61]	[-.14, .03]	[-.34, -.19]	[-.09, .08]	[-.11, .05]	[-.19, -.03]	
10. sem_recall_count	0.88	1.34	.16**	.09*	.39**	.03	-.11*	.07	.04	-.11*	.48**
			[.08, .24]	[.00, .17]	[.32, .46]	[-.05, .12]	[-.19, -.02]	[-.02, .15]	[-.04, .12]	[-.19, -.02]	[.41, .54]

11. semantic_recognition	0.74	0.44	.17**	.19**	.18**	.04	-.10*	.05	.01	-.03	.27**	.25**
			[.09, .25]	[.11, .27]	[.10, .26]	[-.04, .12]	[-.18, -.01]	[-.04, .13]	[-.08, .09]	[-.11, .05]	[.19, .35]	[.17, .32]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .001$.

Table S 10 *Effects of Individual Lexical Knowledge (Vocabulary Checklist) on Caregivers' Verbal Scaffolds Expressed in Incidence Rate Ratios*

	Verbal scaffolds							
Predictors	*Incidence Rate Ratios*	*SE*	*CI*	*p*	*Incidence Rate Ratios*	*SE*	*CI*	*p*
(Intercept)	1.13	0.40	0.56 – 2.28	0.726	1.69	0.65	0.80 – 3.59	0.169
Vocabulary knowledge	1.01	0.03	0.95 – 1.07	0.700	1.01	0.03	0.96 – 1.07	0.672
Age of acquisition	1.07	0.02	1.03 – 1.11	**<0.001**	1.04	0.02	0.99 – 1.08	0.099
Book format [digital]	0.87	0.05	0.79 – 0.97	**0.009**	0.88	0.05	0.80 – 0.98	**0.017**
Individual lexical knowledge					0.82	0.05	0.72 – 0.94	**0.003**
Random Effects								
σ^2	0.36				0.36			
τ_{00}	0.42 $_{child\_ID}$				0.43 $_{child\_ID}$			
ICC	0.54				0.54			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.021 / 0.548				0.029 / 0.555			

Table S 11 *Effects of Individual Lexical Knowledge (Vocabulary Checklist) on Caregivers' Gestural Scaffolds Expressed in Odds Ratio*

	Gestural scaffolds							
Predictors	*Odds Ratios*	*SE*	*CI*	*p*	*Odds Ratios*	*SE*	*CI*	*p*
(Intercept)	9.37	10.75	0.99 – 88.86	0.051	5.32	6.93	0.41 – 68.44	0.200
Vocabulary knowledge	0.99	0.08	0.84 – 1.16	0.882	0.99	0.08	0.84 – 1.16	0.870
Age of acquisition	1.00	0.08	0.86 – 1.16	0.955	1.04	0.10	0.87 – 1.25	0.643
Book format [digital]	0.31	0.08	0.19 – 0.50	**<0.001**	0.30	0.08	0.19 – 0.49	**<0.001**
Individual lexical knowledge					1.31	0.39	0.73 – 2.35	0.359
Random Effects								
σ^2	3.29				3.29			
τ_{00}	2.52 $_{child\_ID}$				2.54 $_{child\_ID}$			
ICC	0.43				0.44			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.056 / 0.465				0.058 / 0.468			

Table S 12 *Effects of Fine-Tuning on Phonological Recall Expressed in Odds Ratios*

	Phonological recall							
Predictors	*Odds Ratios*	*SE*	*CI*	*p*	*Odds Ratios*	*SE*	*CI*	*p*
(Intercept)	0.16	0.07	0.07 – 0.35	**<0.001**	0.02	0.01	0.00 – 0.06	**<0.001**
Vocabulary knowledge	1.09	0.04	1.01 – 1.17	**0.031**	1.11	0.05	1.02 – 1.22	**0.014**
Individual lexical knowledge					21.49	11.08	7.83 – 59.02	**<0.001**
Verbal scaffolds					0.87	0.13	0.65 – 1.18	0.382
Individual lexical knowledge × Verbal scaffolds					1.09	0.18	0.79 – 1.50	0.597
Random Effects								
σ^2	3.29				3.29			
τ_{00}	0.00 $_{child\_ID}$				0.00 $_{child\_ID}$			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.014 / NA				0.473 / NA			

Table S 13 *Effects of Fine-Tuning on Semantic Recall Expressed in Incidence Rate Ratios*

	Semantic recall							
Predictors	*Incidence Rate Ratios*	*SE*	*CI*	*p*	*Incidence Rate Ratios*	*SE*	*CI*	*p*
(Intercept)	0.18	0.07	0.08 – 0.41	**<0.001**	0.06	0.03	0.03 – 0.14	**<0.001**
Vocabulary knowledge	1.13	0.04	1.05 – 1.22	**0.001**	1.13	0.04	1.05 – 1.21	**0.001**
Individual lexical knowledge					5.02	0.86	3.59 – 7.02	**<0.001**
Verbal scaffolds					1.07	0.03	1.01 – 1.13	**0.033**
Individual lexical knowledge × Verbal scaffolds					0.92	0.03	0.86 – 0.99	**0.024**
Random Effects								
σ^2	0.93				0.93			
τ_{00}	0.48 $_{child\_ID}$				0.43 $_{child\_ID}$			
ICC	0.34				0.32			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.074 / 0.392				0.309 / 0.528			

Table S 14 *Effects of Fine-Tuning on Semantic Recognition Expressed in Odds Ratios*

	Semantic recognition							
Predictors	*Odds Ratios*	*SE*	*CI*	*p*	*Odds Ratios*	*SE*	*CI*	*p*
(Intercept)	0.59	0.32	0.20 – 1.68	0.321	0.29	0.18	0.09 – 0.96	**0.043**
Vocabulary knowledge	1.19	0.06	1.07 – 1.31	**0.001**	1.20	0.07	1.07 – 1.34	**0.002**
Individual lexical knowledge					3.01	1.01	1.57 – 5.80	**0.001**
Verbal scaffolds					1.02	0.06	0.91 – 1.14	0.716
Individual lexical knowledge × Verbal scaffolds					1.07	0.10	0.88 – 1.30	0.484
Random Effects								
σ^2	3.29				3.29			
τ_{00}	0.58 $_{child\_ID}$				0.72 $_{child\_ID}$			
ICC	0.15				0.18			
N	78 $_{child\_ID}$				78 $_{child\_ID}$			
Observations	550				550			
Marginal R^2 / Conditional R^2	0.051 / 0.194				0.143 / 0.298			

Supplementary Materials for Chapter 4

Table S 15 *Results of the Model Estimating the Effects of Literacy Skills, Grade, Gender and Language Status on the Probability of Using the Reveal Word Scaffold, Expressed in Odds Ratios*

	Use of the Scaffold			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	0.47	0.03	0.41 – 0.53	**<0.001**
DIBELS [Well Below Benchmark]	1.27	0.01	1.25 – 1.29	**<0.001**
DIBELS [Below Benchmark]	1.25	0.01	1.23 – 1.27	**<0.001**
DIBELS [Above Benchmark]	0.73	0.00	0.72 – 0.74	**<0.001**
Grade [1]	0.57	0.00	0.56 – 0.58	**<0.001**
Grade [2]	0.30	0.00	0.29 – 0.30	**<0.001**
Grade [3]	0.24	0.00	0.23 – 0.24	**<0.001**
Gender [Male]	0.82	0.00	0.81 – 0.83	**<0.001**
Language Status [Dual Language Learner]	1.27	0.01	1.25 – 1.28	**<0.001**
Random Effects				
σ^2	3.29			
$\tau_{00\ q\_id}$	0.12			
$\tau_{00\ level\_name}$	0.37			
ICC	0.13			
$N_{\ q\_id}$	505			
$N_{\ level\_name}$	54			
Observations	1716787			
Marginal R^2 / Conditional R^2	0.065 / 0.186			

Table S 16 *Results of the Model Estimating the Effects of the Scaffold, Literacy Skills, Grade, Gender and Language Status on Task Accuracy*

	Accuracy in the Task			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	3.81	0.15	3.53 – 4.12	**<0.001**
Use of Scaffold [yes]	1.30	0.01	1.28 – 1.31	**<0.001**
DIBELS [Well Below Benchmark]	0.68	0.00	0.67 – 0.68	**<0.001**
DIBELS [Below Benchmark]	0.90	0.01	0.89 – 0.92	**<0.001**
DIBELS [Above Benchmark]	1.26	0.01	1.25 – 1.28	**<0.001**
Grade [1]	0.99	0.01	0.97 – 1.00	0.050
Grade [2]	0.97	0.01	0.95 – 0.98	**<0.001**
Grade [3]	1.17	0.01	1.15 – 1.19	**<0.001**
Gender [Male]	0.91	0.00	0.91 – 0.92	**<0.001**
Language Status [Dual Language Learner]	0.81	0.00	0.80 – 0.81	**<0.001**
Random Effects				
σ^2	3.29			
$\tau_{00\ q\_id}$	0.70			
$\tau_{00\ level\_name}$	0.04			
ICC	0.18			
$N_{\ q\_id}$	505			
$N_{\ level\_name}$	54			
Observations	1716787			
Marginal R^2 / Conditional R^2	0.020 / 0.200			

Table S 17 *Results of the Model Estimating the Effects of the Scaffolded Word Type, and Interaction Between Scaffold Use and Literacy Skills, Grade, Gender, and Language Status on the Accuracy in the Task*

	Accuracy			
Predictors	*Estimate*	*SE*	*CI*	*p*
Intercept	3.67	0.16	3.38 – 3.99	**<0.001**
Scaffolded Nouns	1.03	0.01	1.01 – 1.05	**0.001**
Scaffolded Verbs	0.98	0.01	0.96 – 1.00	0.057
Scaffolded Adjectives	0.70	0.02	0.67 – 0.74	**<0.001**
Scaffolded Adverbs	0.84	0.04	0.76 – 0.93	**0.001**
Scaffolded Prepositions	0.98	0.06	0.87 – 1.10	0.712
Use of Scaffold [yes]	1.57	0.03	1.51 – 1.64	**<0.001**
DIBELS[Well Below Benchmark]	0.66	0.00	0.65 – 0.67	**<0.001**
DIBELS [Below Benchmark]	0.89	0.01	0.88 – 0.90	**<0.001**
DIBELS [Above Benchmark]	1.29	0.01	1.27 – 1.30	**<0.001**
Grade [1]	1.03	0.01	1.02 – 1.05	**<0.001**
Grade [2]	1.02	0.01	1.01 – 1.04	**0.009**
Grade [3]	1.25	0.01	1.22 – 1.27	**<0.001**
Gender [Male]	0.91	0.00	0.90 – 0.92	**<0.001**
Language Status [Dual Language Learner]	0.80	0.00	0.80 – 0.81	**<0.001**
Use of Scaffold [yes]:DIBELS [Well Below Benchmark]	1.22	0.02	1.19 – 1.26	**<0.001**
Use of Scaffold [yes]:DIBELS [Below Benchmark]	1.11	0.02	1.08 – 1.15	**<0.001**
Use of Scaffold [yes]:DIBELS [Above Benchmark]	0.86	0.01	0.83 – 0.88	**<0.001**
Use of Scaffold [yes]:Grade [1]	0.80	0.01	0.77 – 0.82	**<0.001**
Use of Scaffold [yes]:Grade [2]	0.71	0.01	0.69 – 0.73	**<0.001**

Use of Scaffold [yes]:Grade [3]	0.67	0.01	0.64 – 0.70	**<0.001**
Use of Scaffold [yes]:Gender [Male]	1.03	0.01	1.01 – 1.05	**0.015**
Use of Scaffold [yes]:Language Status [Bilingual]	1.03	0.01	1.00 – 1.05	**0.023**

Random Effects

σ^2	3.29
$\tau_{00\ q\_id}$	0.69
$\tau_{00\ level\_name}$	0.05
ICC	0.18
$N_{\ q\_id}$	505
$N_{\ level\_name}$	54
Observations	1716787
Marginal R^2 / Conditional R^2	0.021 / 0.201

Consolidated Bibliography

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*(3), 199–226. https://doi.org/10.1080/09588220802090246

Abu Rabia, S. (2019). The effect of degrees of bilingualism on metacognitive linguistic skills. *International Journal of Bilingualism*, *23*(5), 1064–1086. https://doi.org/10.1177/1367006918781060

Adlof, S. M., Baron, L. S., Bell, B. A., & Scoggins, J. (2021). Spoken Word Learning in Children With Developmental Language Disorder or Dyslexia. *Journal of Speech, Language, and Hearing Research*, *64*, 2734–2749. https://doi.org/10.23641/asha

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of english as a foreign language. *Instructional Science*, *35*(6), 499–517. https://doi.org/10.1007/s11251-007-9016-7

Allen, M. L., Hartley, C., & Cain, K. (2016). iPads and the use of "apps" by children with autism spectrum disorder: Do they promote learning? *Frontiers in Psychology*, *7*(AUG), 1–7. https://doi.org/10.3389/fpsyg.2016.01305

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization [version 1; peer review: 2 approved]. *Wellcome Open Research*, *4*, 1–40. https://doi.org/10.12688/wellcomeopenres.15191.1

Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning Foreign Language Vocabulary with Gestures and Pictures Enhances Vocabulary Memory for Several Months Post-Learning in Eight-Year-Old School Children. *Educational Psychology Review*, *32*(3), 815–850. https://doi.org/10.1007/s10648-020-09527-z

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary Development: A Morphological Analysis. *Monographs of the Society for Research in Child Development*, *58*(10), 1–186.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, *105*(1), 158–173.

Baker, L., & Cerro, L. C. (2000). *Assessing Metacognition in Children and Adults Assessing Metacognition in Children and Adults*.

Barnes, E. M., Hadley, E. B., Lawson-Adams, J., & Dickinson, D. K. (2023). Nonverbal supports for word learning: Prekindergarten teachers' gesturing practices during shared book reading. *Early Childhood Research Quarterly*, *64*, 302–312. https://doi.org/10.1016/j.ecresq.2023.04.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergman Deitcher, D., Johnson, H., & Aram, D. (2019). Does book genre matter? Boys' and girls' word learning from narrative and informational books in the preschool years. *Journal of Research in Reading*, *42*(1), 193–211. https://doi.org/10.1111/1467-9817.12266

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology*, *150*, 404–410. https://doi.org/10.1016/j.jecp.2016.06.009

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared Book Reading: When and How Questions Affect Young Children's Word Learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Blything, L. P., Hardie, A., & Cain, K. (2020). Question Asking During Reading Comprehension Instruction: A Corpus Study of How Question Type Influences the Linguistic Complexity of Primary School Students' Responses. *Reading Research Quarterly*, *55*(3), 443–472. https://doi.org/10.1002/rrq.279

Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Pearson, R. M. (2018). Stability of core language skill from infancy to adolescence in typical and atypical development. *Science Advances*, *4*(11). https://doi.org/10.1126/sciadv.aat7422

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*, *19*(6), 918–932. https://doi.org/10.1111/desc.12343

Brinchmann, E. I., Røe-Indregård, H., Karlsen, J., Schauber, S. K., & Hagtvet, B. E. (2023). The linguistic complexity of adult and child contextualized and decontextualized talk. *Child Development*, *94*(5), 1368–1380. https://doi.org/10.1111/cdev.13932

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand English word lemmas. *Behavior Research Methods*, *46*(3), 904–911.

Bus, A. G., Neuman, S. B., & Roskos, K. (2020). Screens, Apps, and Digital Books for Young Children: The Promise of Multimedia. *AERA Open*, *6*(1), 233285842090149. https://doi.org/10.1177/2332858420901494

Cain, K. (2007). Deriving word meanings from context: Does explanation facilitate contextual analysis? *Journal of Research in Reading*, *30*(4), 347–359. https://doi.org/10.1111/j.1467-9817.2007.00336.x

Cain, K., Lemmon, K., & Oakhill, J. (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. https://doi.org/10.1037/0022-0663.96.4.671

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), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Cain, K., & Oakhill, J. V. (2018). Vocabulary development and reading comprehension: a reciprocal relationship. *Oxford University Press.*

Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension

difficulties. *Journal of Child Language*, *30*(3), 681–694. https://doi.org/10.1017/S0305000903005713

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-Based Learning in Higher Education: A Meta-Analysis. *Review of Educational Research*, *90*(4), 499–541. https://doi.org/10.3102/0034654320933544

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978. https://doi.org/10.1037/0012-1649.38.6.967

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *4*, 284–290.

Clark, C. (2011). *Boys, girls and communication: Their views, confidence and why these skills matter*. London: National Literacy Trust.

Clark, K. F., & Graves, M. F. (2005). Scaffolding Students' Comprehension of Text. *The Reading Teacher*, *58*(6), 570–580. https://doi.org/10.1598/rt.58.6.6

Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal*, *19*(2), 237–248. https://doi.org/10.3102/00028312019002237

Language and Reading Research Consortium, Currie, N. K., & Muijselaar, M. M. L. (2019). Inference Making in Young Children: The Concurrent and Longitudinal Contributions of Verbal Working Memory and Vocabulary. *Journal of Educational Psychology*, *111*(8), 1416–1431. https://doi.org/10.1037/edu0000342

Courage, M. L. (2019). *From Print to Digital: The Medium Is Only Part of the Message* (pp. 23–43). https://doi.org/10.1007/978-3-030-20077-0_3

Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in

computer-based training environments. *Instructional Science*, *30*(6), 433–464. https://doi.org/10.1023/A:1020516301541

Cunningham, A. E., & Stanovich, K. E. (1998). What Reading Does for the Mind. *Journal of Direct Instruction*, *1*(2), 137–149.

Dalton, B., Proctor, C. P., Uccelli, P., Mo, E., & Snow, C. E. (2011). Designing for diversity: The role of reading strategies and interactive vocabulary in a digital reading environment for fifth-grade monolingual English and Bilingual students. *Journal of Literacy Research*, *43*(1), 68–100. https://doi.org/10.1177/1086296X10397872

Dawson, N., Hsiao, Y., Wei, A., Tan, M., & Banerji, N. (2021). Features of lexical richness in children's books: Comparisons with child-directed speech. *Language Development Research*, *1*(1), 9–53. https://doi.org/10.34842/5we1-yk94

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica*, *144*(1), 173–179. https://doi.org/10.1016/j.actpsy.2013.05.013

de Nooijer, J. A., van Gog, T., Paas, F., & Zwaan, R. A. (2014). Words in action. *Gesture*, *14*(1), 46–69. https://doi.org/10.1075/gest.14.1.03noo

De Sousa, I., & Oakhill, J. (1996). Do levels of interest have an effect on children's comprehension monitoring performance? *British Journal of Educational Psychology*, *66*(4), 471–482. https://doi.org/10.1111/j.2044-8279.1996.tb01213.x

Degener, S., & Berne, J. (2017). Complex Questions Promote Complex Thinking. *Reading Teacher*, *70*(5), 595–599. https://doi.org/10.1002/trtr.1535

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, *25*, 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

Diprossimo, L., Ushakova, A., Zoski, J., Gamble, H., Irey, R., & Cain, K. (2023). The associations between child and item characteristics, use of vocabulary scaffolds, and reading comprehension in a digital environment: Insights from a big data

approach. *Contemporary Educational Psychology*, *73*. https://doi.org/10.1016/j.cedpsych.2023.102165

Donnellan, E., Jordan-Barros, A., Theofilogiannakou, N., Brekelmans, G., Murgiano, M., Motamedi, Y., Grzyb, B., Gu, Y., & Vigliocco, G. (2023). The impact of caregivers' multimodal behaviours on children's word learning: A corpus-based investigation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 693–701.

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis. *Child Development*, *91*(2), e383–e399. https://doi.org/10.1111/cdev.13225

Duncan, L. G., Mcgeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology*, *107*(2), 209–238. https://doi.org/10.1111/bjop.12134

Dunn, L. M., & Dunn, D. M. (2009). The British picture vocabulary scale. In *GL Assessment Limited.*

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. https://doi.org/10.1080/10888438.2013.774005

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

Elley, W. B. (1989). Vocabulary Acquisition from Listening to Stories. *Reading Research Quarterly*, *24*(2), 174–187.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, *43*(4), 559–617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Eng, C. M., Tomasic, A. S., & Thiessen, E. D. (2019). Contingent Responsivity in E-Books Modeled From Quality Adult-Child Interactions: Effects on Children's Learning and

Attention. *Developmental Psychology*, *56*(2), 285–297. https://doi.org/10.1037/dev0000869

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Gayraud, F., Kovacevic, M., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology*, *30*(2), 326–343. https://doi.org/10.1111/j.2044-835X.2011.02042.x

Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, *32*(3), 343–364. https://doi.org/10.1177/0142723711422626

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). The MacArthur Communicative Development Inventories: User's guide and technical manual. In *San Diego, CA: Singular Publishing Group.*

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). *Variability in Early Communicative Development*. *59*(5).

Fernald, A., & Mazzie, C. (1991). Prosody and Focus in Speech to Infants and Adults. *Developmental Psychology*, *27*(2), 209–221. https://doi.org/10.1037/0012-1649.27.2.209

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R* (Sage publications).

Fisch, S. M. (2000). A capacity model of children's comprehension of educational content on television. *Media psychology*, *2*(1), 63-91. https://doi.org/10.1207/S1532785XMEP0201_4

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The Effects of Shared Storybook Reading on Word Learning: A Meta-Analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/http://dx.doi.org/10.1037/dev0000512

Flack, Z. M., & Horst, J. S. (2018). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development*, *27*(1), 1–12. https://doi.org/10.1002/icd.2047

Flewitt, R., Messer, D., & Kucirkova, N. (2015). New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy*, *15*(3), 289–310. https://doi.org/10.1177/1468798414533560

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Francey, G., & Cain, K. (2015). Effect of imagery training on children's comprehension of pronouns. *Journal of Educational Research*, *108*(1), 1–9. https://doi.org/10.1080/00220671.2013.824869

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). Variability and Consistency in Early Language Learning. In *Variability and Consistency in Early Language Learning*. The MIT Press. https://doi.org/10.7551/mitpress/11577.001.0001

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A Comparison of Children's Reading on Paper Versus Screen: A Meta-Analysis. *Review of Educational Research*, *91*(4), 483–517. https://doi.org/10.3102/0034654321998074

Fusaroli, R., Weed, E., Rocca, R., Fein, D., & Naigles, L. (2023). Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development. *Cognition*, *236*. https://doi.org/10.1016/j.cognition.2023.105422

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological Short-Term Memory and New Word Learning in Children. *Developmental Psychology*, *33*(6), 966–979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological Short-term Memory and Vocabulary Development: Further Evidence on the Nature of the Relationship. *Applied Cognitive Psychology*, *13*(1), 65–77. https://doi.org/10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O

Gentner, D. (1982). Why nouns are learned before verbs. *Language Development Vol. 2*, 301–334.

Gioia, G. A. , Espy, K. A. , & Isquith, P. K. (2003). BRIEF-P: Behavior rating inventory of executive function - preschool version. In *Psychological Assessment Resources (PAR).*

Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, *4*, 189–202.

Gonzalez, M. (2014). The effect of embedded text-to-speech and vocabulary ebook scaffolds on the comprehension of students with reading disabilities. *International Journal of Special Education*, *29*(3), 111–125.

Government of the UK. (2019). *English indices of deprivation 2019*. https://imd-by-postcode.opendatacommunities.org/imd/2019

Government of the UK. (2022, December). *Regional ethnic diversity*. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest/

Gracia, P., Garcia-Roman, J., Oinas, T., & Anttila, T. (2020). Child and Adolescent Time Use: A Cross-National Study. *Journal of Marriage and Family*, *82*(4), 1304–1325. https://doi.org/10.1111/jomf.12626

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, *18*(5), 797–824. https://doi.org/10.1111/infa.12006

Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Kim, E., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Journal of Applied Developmental Psychology*, *77*, 101332. https://doi.org/10.1016/j.appdev.2021.101332

Hadley, E. B., & Dickinson, D. K. (2019). Cues for word-learning during shared book-reading and guided play in preschool. *Journal of Child Language*, *46*(6), 1202–1227. https://doi.org/10.1017/S0305000919000552

Hadley, E. B., & Dickinson, D. K. (2020). Measuring young children's word knowledge: A conceptual review. *Journal of Early Childhood Literacy*, *20*(2), 223–251. https://doi.org/10.1177/1468798417753713

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Building Semantic Networks: The Impact of a Vocabulary Intervention on Preschoolers' Depth of Word Knowledge. *Reading Research Quarterly*, *54*(1), 41–61. https://doi.org/10.1002/rrq.225

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181–198. https://doi.org/10.1002/rrq.130

Hald, L. A., de Nooijer, J., van Gog, T., & Bekkering, H. (2016). Optimizing Word Learning via Links to Perceptual and Motoric Experience. *Educational Psychology Review*, *28*(3), 495–522. https://doi.org/10.1007/s10648-015-9334-2

Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*, *37*(2), 205–225. https://doi.org/10.1177/0142723716679956

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hartley, C., Bird, L. A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*, 104265. https://doi.org/10.1016/j.cognition.2020.104265

Hassinger-Das, B., Brennan, S., Dore, R. A., Golinkoff, R. M., & Hirsh-Pasek, K. (2020). Children and Screens. *Annual Review of Developmental Psychology*, *2*(1), 69–92. https://doi.org/10.1146/annurev-devpsych-060320-095612

Hills, T. (2013). The company that words keep: Comparing the statistical structure of child- Versus adult-Directed language. *Journal of Child Language*, *40*(3), 586–604. https://doi.org/10.1017/S0305000912000165

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4–14. https://doi.org/10.1037/a0027238

Honma, M., Masaoka, Y., Iizuka, N., Wada, S., Kamimura, S., Yoshikawa, A., Moriya, R., Kamijo, S., & Izumizaki, M. (2022). Reading on a smartphone affects sigh generation, brain activity, and comprehension. *Scientific Reports*, *12*(1), 1–8. https://doi.org/10.1038/s41598-022-05605-0

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

James, E., Currie, N. K., Tong, S. X., & Cain, K. (2021). The relations between morphological awareness and reading comprehension in beginner readers to young adolescents. *Journal of Research in Reading*, *44*(1), 110–130. https://doi.org/10.1111/1467-9817.12316

Joffe, V. L., Cain, K., & Marić, N. (2007). Comprehension problems in children with specific language impairment: Does mental imagery training help? *International Journal of Language and Communication Disorders*, *42*(6), 648–664. https://doi.org/10.1080/13682820601084402

Kaiser, H. F. (1974). An index of factorial simplicity. *psychometrika*, *39*(1), 31-36.

Kapa, L. L., & Erikson, J. A. (2020). The relationship between word learning and executive function in preschoolers with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *63*(7), 2293–2307. https://doi.org/10.1044/2020_JSLHR-19-00342

Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences*, *27*, 82–89. https://doi.org/10.1016/j.lindif.2013.07.002

Kihlstrom, J. F. (2021). Ecological Validity and "Ecological Validity." *Perspectives on Psychological Science*, *16*(2), 466–471. https://doi.org/10.1177/1745691620966791

Kim, D., & Gilman, D. A. (2008). Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning. *Journal of Educational Technology & Society*, *11*(3), 114–126.

Kim, N. J., Belland, B. R., & Walker, A. E. (2018). Effectiveness of Computer-Based Scaffolding in the Context of Problem-Based Learning for Stem Education: Bayesian Meta-analysis. *Educational Psychology Review*, *30*(2), 397–429. https://doi.org/10.1007/s10648-017-9419-1

Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., & Gruelich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing*, *27*(2), 237–253. https://doi.org/10.1007/s11145-013-9440-9

Kim, Y.-S. (2020). Hierarchical and Dynamic Relations of Language and Cognitive Skills to Reading Comprehension: Testing the Direct and Indirect Effects Model of Reading (DIER). *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000407.supp

Kirby, J. R., & Moore, P. J. (1987). Metacognitive awareness about reading and its relation to reading ability. *Journal of Psychoeducational Assessment*, *2*, 119–137.

Korat, O., & Or, T. (2010). How new technology influences parent-child interaction: The case of e-book reading. *First Language*, *30*(2), 139–154. https://doi.org/10.1177/0142723709359242

Korat, O., & Shamir, A. (2012). Direct and indirect teaching: Using e-books for supporting vocabulary, word reading, and story comprehension for young children. *Journal of Educational Computing Research*, *46*(2), 135–152. https://doi.org/10.2190/EC.46.2.b

Korkman, M., Kirk, U., & Kemp, S. (1998). NESPY: A developmental neuropsychological assessment. In *San Antonio, TX: The Psychological Corporation.*

Kucirkova, N. (2017). An integrative framework for studying, designing and conceptualising interactivity in children's digital books. *British Educational Research Journal*, *43*(6), 1168–1185. https://doi.org/10.1002/berj.3317

Kucirkova, N. (2019). Children's Reading With Digital Books: Past Moving Quickly to the Future. *Child Development Perspectives*, *13*(4), 208–214. https://doi.org/10.1111/cdep.12339

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*(4), 978–990.

Lämsä, J., Hämäläinen, R., Aro, M., Koskimaa, R., & Äyrämö, S. M. (2018). Games for enhancing basic reading and maths skills: A systematic review of educational game design in supporting learning by people with learning disabilities. *British Journal of Educational Technology*, *49*(4), 596–607. https://doi.org/10.1111/bjet.12639

LARRC. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, *50*(2), 151–169. https://doi.org/10.1002/rrq.99

Laurence, B. L., Patricia, D., Jeffrey, D. K., Sharon, L. C., & Justin, B. K. (2020). After Initial Retrieval Practice, More Retrieval Produces Better Retention Than More Study in the Word Learning of Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *63*, 2763–2776.

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes. *Journal of Speech, Language, and Hearing Research*, *52*, 1401–1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2019). The effects of questions during shared-reading: Do demand-level and placement really matter? *Early Childhood Research Quarterly*, *47*, 49–61. https://doi.org/10.1016/j.ecresq.2018.10.006

Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2020). More than words: Narrator engagement during storytelling increases children's word learning, story comprehension, and on-task behavior. *Early Childhood Research Quarterly*, *51*, 338–351. https://doi.org/10.1016/j.ecresq.2019.12.009

Lerner, M. D., & Lonigan, C. J. (2014). Executive Function Among Preschool Children: Unitary Versus Distinct Abilities. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 626–639. https://doi.org/10.1007/s10862-014-9424-3

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. In *American Educational Research Journal* (Vol. 47, Issue 3). https://doi.org/10.3102/0002831209355469

Leung, A., Tunkel, A., & Yurovsky, D. (2021). Parents Fine-Tune Their Speech to Children's Vocabulary Knowledge. *Psychological Science*, *32*(7), 975–984. https://doi.org/10.1177/0956797621993104

Liu, Y. T., & Leveridge, A. N. (2017). Enhancing L2 vocabulary acquisition through implicit reading support cues in e-books. *British Journal of Educational Technology*, *48*(1), 43–56. https://doi.org/10.1111/bjet.12329

Logan, J. A. R., Justice, L. M., Yumus¸ Melike, & Chaparro-Moreno, L. J. (2019). When Children Are Not Read to at Home: The Million Word Gap. *Journal of Developmental Behavioral Pediatrics*, *40*(5), 383–386. DOI: 10.1097/DBP.0000000000000657

Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, *62*(2), 175–187. https://doi.org/10.1080/00131911003637006

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant-and adult-directed speech. *Language Learning and Development*, *7*(3), 185–201. https://doi.org/10.1080/15475441.2011.579839

Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R., & Christakis, D. A. (2020). Associations between Screen Use and Child Language Skills: A Systematic Review and Meta-analysis. *JAMA Pediatrics*, *174*(7), 665–675. https://doi.org/10.1001/jamapediatrics.2020.0327

Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M. A., Guerrero-Puerta, L., Aguilar-Parra, J. M., Trigueros, R., & Alias, A. (2021). Between level up and game over: A systematic literature review of gamification in education. *Sustainability*, *13*(4), 1–14. https://doi.org/10.3390/su13042247

Masek, L. R., McMillan, B. T. M., Paterson, S. J., Tamis-LeMonda, C. S., Golinkoff, R. M., & Hirsh-Pasek, K. (2021). Where language meets attention: How contingent interactions promote learning. *Developmental Review*, *60*. https://doi.org/10.1016/j.dr.2021.100961

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/s0959-4752(02)00016-6

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, *33*(5), 403–423. https://doi.org/10.1111/jcal.12197

Mayer, R. E. (2020). Where is the learning in mobile technologies for learning? *Contemporary Educational Psychology*, *60*(December 2019), 101824. https://doi.org/10.1016/j.cedpsych.2019.101824

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mayer, R. E., & Sims, V. K. (1994). For Whom Is a Picture Worth a Thousand Words? Extensions of a Dual-Coding Theory of Multimedia Learning. *American Psychological Association, Inc*, *86*(3), 389–401.

Mckeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking Reading Comprehension. *Reading Research Quaterly*, *44*(3), 218–253.

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*(March 2019), 104092. https://doi.org/10.1016/j.jml.2020.104092

Ministry of Housing Communities and Local Government. (2019). *The English Indices of Deprivation 2019: Research Report*. https://assets.publishing.service.gov.uk/media/5d8b364ced915d03709e3cf2/IoD2019_Research_Report.pdf

Mol, S. E., & Bus, A. G. (2011). To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, *39*(5), 527–546. https://doi.org/10.1177/0142723719849996

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Montroy, J. J., Merz, E. C., Williams, J. M., Landry, S. H., Johnson, U. Y., Zucker, T. A., Assel, M., Taylor, H. B., Lonigan, C. J., Phillips, B. M., Clancy-Menchetti, J., Barnes, M. A., Eisenberg, N., Spinrad, T., Valiente, C., de Villiers, J., & de Villiers, P. (2019). Hot and cool dimensionality of executive function: Model invariance across age and maternal education in preschool children. *Early Childhood Research Quarterly*, *49*, 188–201. https://doi.org/10.1016/j.ecresq.2019.06.011

Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. (2019). Differences in parent-toddler interactions with electronic versus print books. *Pediatrics*, *143*(4). https://doi.org/10.1542/peds.2018-2012

Nassaji, H. (2003). L2 Vocabulary Learning From Context: Strategies, Knowledge Sources, and Their Relationship With Success in L2 Lexical Inferencing. *Tesol Quarterly*, *37*(4), 645–670.

Nation, K., Dawson, N. J., & Hsiao, Y. (2022). Book Language and Its Implications for Children's Language, Literacy, and Development. In *Current Directions in Psychological Science* (Vol. 31, Issue 5, p. 464). SAGE Publications Inc. https://doi.org/10.1177/09637214221119448

Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*(1), 43–79. https://doi.org/10.1017/S0305000900005584

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*(October 2018), 100290. https://doi.org/10.1016/j.edurev.2019.100290

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. https://doi.org/10.1080/10888438.2010.529219

Ofcom. (2019). *Children's and parents' media use and attitudes*. https://www.ofcom.org.uk/__data/assets/pdf_file/0027/134892/Children-and-Parents-Media-Use-and-Attitudes-Annex-1.pdf

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*(3), 423–430. https://doi.org/10.1016/j.cognition.2008.10.004

O'Rear, C. D., Seip, I., Azar, J., Baroody, A. J., & McNeil, N. M. (2023). Features in children's counting books that lead dyads to both count and label sets during shared book reading. *Child Development*, *94*(4), 985–1001. https://doi.org/10.1111/cdev.13915

O'Toole, K. J., & Kannass, K. N. (2018). Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology*, *173*, 100–115. https://doi.org/10.1016/j.jecp.2018.03.013

Ozturk, G., & Hill, S. (2020). Mother–child interactions during shared reading with digital and print books. *Early Child Development and Care*, *190*(9), 1425–1440. https://doi.org/10.1080/03004430.2018.1538977

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*, 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Pesco, D., & Gagné, A. (2017). Scaffolding Narrative Skills: A Meta-Analysis of Instruction in Early Childhood Settings. *Early Education and Development*, *28*(7), 773–793. https://doi.org/10.1080/10409289.2015.1060800

Price, J., & Kalil, A. (2019). The Effect of Mother–Child Reading Time on Children's Reading Skills: Evidence From Natural Within-Family Variation. *Child Development*, *90*(6), e688–e702. https://doi.org/10.1111/cdev.13137

Proctor, C. P., Dalton, B., & Grisham, D. L. (2007). Scaffolding English language learners and struggling readers in a universal literacy environment with embedded strategy instruction and vocabulary support. *Journal of Literacy Research*, *39*(1), 71–93. https://doi.org/10.1080/10862960709336758

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental Relations Between Vocabulary Knowledge and Reading Comprehension: A Latent Change Score Modeling Study. *Child Development*, *86*(1), 159–175. https://doi.org/10.1111/cdev.12292

Reese, E., Cox, A., Harte, D., & McAnally, H. (2003). Diversity in adults' styles of reading books to children. *On Reading Books to Children: Parents and Teachers*, *2003*, 35–54. https://doi.org/10.4324/9781410607355

Reich, S. M., Yau, J. C., Xu, Y., Muskat, T., Uvalle, J., & Cannata, D. (2019). Digital or Print? A Comparison of Preschoolers' Comprehension, Vocabulary, and Engagement From a Print Book and an e-Book. *AERA Open*, *5*(3), 233285841987838. https://doi.org/10.1177/2332858419878389

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, *7*(3), 219-237.

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. *Northwestern University, Evanston, Illinois.*

Richter, A., & Courage, M. L. (2017). Comparing electronic and paper storybooks for preschoolers: Attention, engagement, and recall. *Journal of Applied*

Developmental Psychology, *48*, 92–102. https://doi.org/10.1016/j.appdev.2017.01.002

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The Pace of Vocabulary Growth Helps Predict Later Vocabulary Skill. *Child Development*, *83*(2), 508–525. https://doi.org/10.1111/j.1467-8624.2011.01710.x

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, *38*(2), 109–117. https://doi.org/10.1016/j.cedpsych.2012.12.001

Rowland, C. (2013). Understanding child language acquisition. In *Understanding Child Language Acquisition*. https://doi.org/10.4324/9780203776025

Sadeghi, K., & Khezrlou, S. (2012). Glossing mode in self-regulated vocabulary learning, and its relationship with gender, age, and field of study. *Journal of Asia TEFL*, *9*(3), 51–74.

Sadoski, M., & Lawrence, B. (2023). Abstract Vocabulary Development: Embodied Theory and Practice. In *Educational Psychology Review* (Vol. 35, Issue 3). Springer. https://doi.org/10.1007/s10648-023-09802-9

Salmerón, L., Altamura, L., Delgado, P., Karagiorgi, A., & Vargas, C. (2023). Reading comprehension on handheld devices versus on paper: A narrative review and meta-analysis of the medium effect and its moderators. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000830

Salmerón, L., Delgado, P., Vargas, C., & Gil, L. (2021). Tablets for all? Testing the screen inferiority effect with upper primary school students. *Learning and Individual Differences*, *86*. https://doi.org/10.1016/j.lindif.2021.101975

Salmerón, L., & García, V. (2011). Reading skills and children's navigation strategies in hypertext. *Computers in Human Behavior*, *27*(3), 1143–1151. https://doi.org/10.1016/j.chb.2010.12.008

Salomon, G. (1984). Television is "easy" and print is "tough": The differential

investment of mental effort in learning as a function of perceptions and attributions. Journal of Educational Psychology, 76, 647–658. https://psycnet.apa.org/doi/10.1037/0022-0663.76.4.647

Sarı, B., Başal, H. A., Takacs, Z. K., & Bus, A. G. (2019). A randomized controlled trial to test efficacy of digital enhancements of storybooks in support of narrative comprehension and word learning. *Journal of Experimental Child Psychology*, *179*, 212–226. https://doi.org/10.1016/j.jecp.2018.11.006

Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children's word learning. *Developmental Psychology*, *52*(6), 879–886. https://doi.org/10.1037/dev0000125

Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, *29*(4), 531–553. https://doi.org/10.1080/10862969709547973

Scionti, N., & Marzocchi, G. M. (2021). The dimensionality of early executive functions in young preschoolers: Comparing unidimensional versus bidimensional models and their ecological validity. *Child Neuropsychology*, *27*(4), 491–515. https://doi.org/10.1080/09297049.2020.1868419

Sénéchal, M., & Cornell, E. H. (1993). Vocabulary Acquisition through Shared Reading Experiences. *Reading Research Quarterly*, *28*(4), 360–374.

Sénéchal, M., Pagan, S., Lever, R., & Ouellette, G. P. (2008). Relations among the frequency of shared reading and 4-year-old children's vocabulary, morphological and syntax comprehension, and narrative skills. *Early Education and Development*, *19*(1), 27–44. https://doi.org/10.1080/10409280701838710

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual Differences in 4-Year-Old Children's Acquisition of Vocabulary During Storybook Reading. *Journal of Educational Psychology*, *87*(2), 218–229. https://doi.org/10.1037/0022-0663.87.2.218

Shi, J., Gu, Y., & Vigliocco, G. (2022). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*. https://doi.org/10.1111/desc.13357

Smith, L., & Gasser, M. (2005). The Development of Embodied Cognition: Six Lessons from Babies. *Artificial Life*, *11*(1–2), 13–29.

Snow, C. E. (1972). Mothers' Speech to Children Learning Language. *Child Development*, *43*(2), 549–565.

Stanovich, K. E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/rrq.21.4.1

Syrett, K., Musolino, J., & Gelman, R. (2012). How Can Syntax Support Number Word Acquisition? *Language Learning and Development*, *8*(2), 146–176. https://doi.org/10.1080/15475441.2011.583900

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and Pitfalls of Multimedia and Interactive Features in Technology-Enhanced Storybooks: A Meta-Analysis. *Review of Educational Research*, *85*(4), 698–739. https://doi.org/10.3102/0034654314566989

Thompson, R. B., Cothran, T., & McCall, D. (2012). Gender and age effects interact in preschoolers' help-seeking: Evidence for differential responses to changes in task difficulty. *Journal of Child Language*, *39*(5), 1107–1120. https://doi.org/10.1017/S030500091100047X

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. In *Harvard University Press*.

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning*, *59*(3), 623–646. https://doi.org/10.1111/j.1467-9922.2009.00519.x

Troseth, G. L., Strouse, G. A., Flores, I., Stuckelman, Z. D., & Russo Johnson, C. (2020). An enhanced eBook facilitates parent–child talk during shared reading by families of low socioeconomic status. *Early Childhood Research Quarterly*, *50*, 45–58. https://doi.org/10.1016/j.ecresq.2019.02.009

Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, *27*(1), 78–102. https://doi.org/10.1093/applin/ami046

Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221-243. https://doi.org/10.1111/jcom.12024

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Verhoeven, L., & Perfetti, C. A. (2011). Introduction to this special issue: Vocabulary growth and reading skill. Scientific Studies of Reading, *15*(1), 1-7.

Vulchanova, M., Baggio, G., Cangelosi, A., & Smith, L. (2017). Language development in the digital age. *Frontiers in Human Neuroscience*, *11*, 447.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Wainwright, B. R., Allen, M. L., & Cain, K. (2020). Narrative comprehension and engagement with e-books vs. paper-books in autism spectrum condition. *Autism and Developmental Language Impairments*, *5*. https://doi.org/10.1177/2396941520917943

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., & et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559. https://doi.org/10.1037//0012-1649.24.4.552

Wiig, E. H. , Secord, W. A. , & Semel, E. (2004). Clinical evaluation of language fundamentals—Preschool, (CELF Preschool-2). In *Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.*

Wolf, M., & Katzir-Cohen, T. (2009). Reading Fluency and Its Intervention. *Scientific Studies of Reading*, *5*(3), 211–239. https://doi.org/10.1207/S1532799XSSR0503

Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wu, J. Y. (2014). Gender differences in online reading engagement, metacognitive strategies, navigation skills and reading literacy. *Journal of Computer Assisted Learning*, *30*(3), 252–271. https://doi.org/10.1111/jcal.12054

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, *24*(1), 39–58. https://doi.org/10.1080/09588221.2010.523285

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning*, *34*(5–6), 751–777. https://doi.org/10.1080/09588221.2019.1640745

Zucker, T. A., Justice, L. M., Piasta, S. B., & Kaderavek, J. N. (2010). Preschool teachers' literal and inferential questions and children's responses during whole-class shared reading. *Early Childhood Research Quarterly*, *25*(1), 65–83. https://doi.org/10.1016/j.ecresq.2009.07.001