Thesis submitted for the award

Doctor in Education (EdD)

Comparing the impact of OpenDyslexic and Arial fonts on the reading performance of Key Stage 2 readers with dyslexia.

Liz Broadbent

UCL Institute of Education

Date: April 2023

Declaration

I, Liz Broadbent, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word count: 44,761

Abstract

Several fonts have been designed with the aim of ameliorating some of the reading difficulties experienced by those with dyslexia. Anecdotal reports assert that the use of the dyslexia-friendly font OpenDyslexic mitigates reading difficulties by enhancing legibility through unique letterforms but there are few methodologically rigorous, peer-reviewed studies to substantiate or refute these claims. Without empirical evidence it may not be prudent for educational professionals to recommend that readers with dyslexia use a specific font.

To investigate the impact of font on reading performance this mixed methods study compared the test scores in reading accuracy, reading rate and reading comprehension of 40 Key Stage 2 (KS2) participants with dyslexia and a control group of 38 typically progressing KS2 readers when texts were presented in the fonts OpenDyslexic and Arial. The spacing effect of the default designs of the two fonts was considered by including an expanded version of Arial. A semi-structured interview enabled all participants to voice their preferences and opinions of the two fonts.

Findings showed that participants in both the groups achieved significantly higher test scores in reading accuracy and reading rate when passages were presented in OpenDyslexic font. No significant effect of font was found on reading comprehension scores for either group. The variable of spacing did not demonstrate a significant impact on test scores recorded. Readers' preferences for font design were

influenced by a number of subjective factors and did not align consistently with reading test scores.

The findings of this study suggest that there may be a benefit to offering all reading matter in OpenDyslexic font to young readers with and without reading difficulties.

Acknowledgements

Firstly I would like to thank Professor Chloë Marshall and Dr Emma Sumner for their help, expertise, support, and time in guiding me through this process.

A huge thank you to all of the participating schools, especially to the hardworking SENCo's, without whom this whole study could not have happened.

Many thanks to the expert and caring staff of the Beatson West of Scotland Cancer Centre.

This study is in loving memory of Dr Chris Broadbent.

Impact statement

The impact of this study can be likened to throwing a small pebble into a pond and watching the ripples travel outwards. Font design and legibility may appear to be a small element of the whole process of learning to read. However, the effect can be like the ripples from the tiny pebble, with the impact being significant for many young readers and the findings disseminating wider than the initial participants.

The first impact was when one participant realised that they could read words more readily and accurately when text was presented in OpenDyslexic font. Their reaction was profound, indicating that the role of font design has immediate, positive effects on some young readers. The Special Needs Coordinator (SENCo) at this child’s school was very interested in this consequence and so decided to incorporate OpenDyslexic into the presentation of resources for certain pupils and to encourage staff to consider font and legibility. I presented some early findings and background information in a staff meeting at that school to illuminate the SENCo’s decision regarding OpenDyslexic font. Other schools in the study were also interested to incorporate OpenDyslexic into their practice. Ripples of impact could easily flow out from this with information travelling informally by word of mouth to other interested parties.

As COVID-19 precluded personal meetings, parents and carers of participants were sent a brief e-mail report of the findings via the school SENCos (appendix 8) to facilitate decisions regarding integration of OpenDyslexic, a free, open-access font, into their child's home reading possibly via e-readers and tablets.

A further ripple effect is through Continuing Professional Development offered by training institutions and local authorities. Certain training providers are incorporating a brief element concerning font design in sessions on inclusion and assistive technology. This shows that the impact of font design is being circulated ever wider. Through my involvement with the reading charity Coram Beanstalk, I was able to present my work at local and regional levels so that reading volunteers could become aware of the role of legibility which may impact their choice of material used with young readers in the many schools that they assist.

In addition to disseminating findings through schools, the results have been presented in university, namely at poster conferences, which enabled interested parties to access the research findings. It is possible that further dissemination of findings may be through future publication or presentations.

Ultimately it is impossible to judge how far the impact could travel. There is reference to my work on the OpenDyslexic website and the abstracts of the posters appear online. A student in the Netherlands enquired about the study so the internet is pushing the impact wider. OpenDyslexic is available in Greek and Cyrillic scripts, so following dissemination at a conference in Greece[1] findings may influence font choice in other alphabetic orthographies and scripts. Tiny ripples may travel far beyond the classrooms where participants read texts in Arial and OpenDyslexic fonts to ascertain if their reading was impacted by font design.

[1] Presented by Professor Chloë Marshall, 30/9/22, University of Thessaloniki, Greece.

Table of Contents

Tables

Figures

Reflective Statement

Introduction

On my first day of the Doctorate in Education (EdD) I felt like the apocryphal traveller, who on seeking the way, is told “I wouldn't start from here”. Everybody on the course had different experiences and ambitions but I felt that there were few in the cohort who had travelled the path that I had. I joined the EdD programme via the Postgraduate Diploma in Social Science Research Methods (PGDip). This meant that I did one taught module of the EdD course and then prior learning from the PGDip allowed me to progress directly to the Institution-Focused Study (IFS) stage, in effect jumping from Year 1 to Year 2 and changing cohorts along the way.

Although the route is certainly academically and scholastically justified, mine was a circuitous path due mainly to the fact that I did not make the decision to apply to join the EdD programme until after I had finished the PGDip. If I had decided earlier then I would have concentrated on making clearer links between the four modules of the PGDip and the elements that were most valuable to mirror the EdD taught modules. Some of my thoughts and reflections may be clouded by the ambiguity of not having decided to follow the EdD programme when doing the PGDip modules. If I had followed the EdD taught modules route I would have already have been on the programme when required to reflect and would have approached the course differently. That being said, I found the PGDip extremely valuable and I doubt that I would have had the skills and knowledge to have been accepted onto the EdD programme without it.

Time for reflection is often in short supply for those in the education sector as there is always pressure to be firefighting the latest demand or crisis. Any consideration of theory and current research is often skimmed over. Often reflection is more about how to avoid disaster and achieve imposed aims immediately, rather than a deep consideration of education theory and its role in current pedagogy.

Post Graduate Diploma in Social Science Research Methods course (PGdip)

This course allowed me to work with a cohort of people with varied backgrounds and specialisms. Most people that I had liaised with before came from the primary sector usually with specialisms in literacy and dyslexia support. Explaining ideas to fellow professionals from different disciplines was invaluable as it helped with the avoidance of jargon, which initially I assumed everyone used and knew. Realising that terms were more specific to my field than I had appreciated meant that I was careful to ensure that fundamental concepts were clear and unambiguous. This was particularly useful in the module concerning designing a research question. Group discussion in this unit showed that sometimes a term can be interpreted differently depending on the discipline in which it is used. This clarified my thinking when progressing through different iterations of the research question.

The short internship placement included in module 2 of the PGDip made the world of research blossom for me and planted the idea that maybe I could progress towards the EdD. I was able to construct and complete a small-scale project using secondary data. This gave practical experience of research and also a deeper understanding of how to link theory and practice.

By trialling a small pilot project in module 3 and developing a research proposal in module 4, I learned that attention to detail and clarity of thought were vital. Reviewing elements of the module also allowed me to develop cohesion of thought and how to improve and refine the proposal by accessing the alternative viewpoint of colleagues.

Foundations of Professionalism (FoP)

Ultimately after much consideration and soul searching I enrolled on the first module of the EdD, Foundations of Professionalism (FoP). This was for me the final taught module and not the first. It was difficult to link this module’s importance and relevance to the PGDip modules and it was not until a previous EdD alumnus gave a presentation to the cohort that its pertinence fell into place. I realised that the impostor syndrome that I had experienced when meeting so many confident and dynamic people was misplaced. The FoP demonstrated that I belong to a profession in which individuals are constantly conditioned into doubting themselves but who should recognise their expertise and experience and see the invaluable ways that this professionalism can be used to support pupils, colleagues and society at large. By developing skills in research I would be able not only to develop my own levels of understanding of my chosen area of expertise but to weave that knowledge into my professional practice. The realisation that an inner conflict between personal ideology and education policy is commonplace enabled me to take a much broader perspective of professional life. This allowed me to move forward. The most important aspect was the revelation that I could move from being a teacher and a consumer of policy to an active researcher looking at how and why theoretical considerations should be recommended. This represented a refocusing of identity.

Institution-Focused Study (IFS)

My IFS served as a pilot study for my thesis. I found the small-scale version of the main project invaluable as it afforded me the opportunity through trial and error and seeking advice to learn how to organise my time and data more efficiently. This skill was definitely needed when handling the larger data set for the thesis. I also realised that almost everything took longer than I had anticipated and that seemingly minor details or adjustments could grow and take unexpected amounts of time to resolve.

The topic of my IFS had developed from my MA dissertation in which the question of font choice as a variable arose when studying the performance of primary pupils reading texts presented as print on paper and on e-readers. This, combined with a comment from a pupil's parent regarding specialist dyslexia fonts led to developing an interest in font design and its impact on the reading performance of young readers. The IFS enabled me to take the idea through a further iteration.

The process of completing the IFS taught me a huge amount. At the time I felt that I was just moving through the study and adapting what I was doing to reach a final draft. In fact, I was honing my skills in writing a literature review without becoming side-tracked by fascinating but irrelevant information; developing problem solving;

and also beginning to understand statistical data analysis, all of which would underpin my thesis and allow me to focus on the greater clarity of thought and analysis required. In retrospect having to complete the IFS before attempting the thesis was probably the most valuable aspect of the EdD when comparing it to other doctoral programmes, as it not only provided a pilot of the methodology but allowed for glitches in my own thought processes and organisation to be ironed out.

Formal review before thesis

The viva for the thesis upgrade was initially daunting but on reflection was an extremely valuable experience. Discussing the proposal allowed me to consider my clarity of thought and the coherence of my writing. The examiners gave an additional perspective to my thought processes. The advice that they gave was extremely useful. They suggested two experimental spacing conditions rather than the three that I had proposed. I adopted this because on reflection trying to compare three spacing conditions would have been challenging due to the larger amount of additional data that would have been generated. Testing the third spacing condition would be valuable in a future study to clarify what was found during this research, but to have included it in this study would have been too cumbersome.

The Thesis stage

The IFS confirmed my main area of study for my thesis and reflection on feedback from my supervisor and second marker allowed me to address defined challenges. This was achieved by attending the Research Training Programme (RTP) summer school to focus on courses in quantitative analysis and writing a literature review. Without the IFS to highlight these weaknesses I may have moved on to the thesis stage without tackling them sufficiently. I also modified the methodology to make the thesis study more focused and coherent.

During the data collection phase, I altered and refined both my research questions and focus. The main alterations were to discard the data from the Key Stage 3 (KS3) participants. The reasons for this were largely twofold. Time restrictions on how long each participant was allowed to be out of class meant that firstly insufficient data could be obtained from each pupil. This was mainly due to the length of the reading passages and the detail required for the York Assessment of Reading for Comprehension (YARC) analysis tool to be meaningful. Secondly, the data obtained using the YARC secondary edition, KS3, could not be compared sufficiently accurately or informatively with the Key Stage 2 data. I felt that a separate study with KS3 participants would be more valid than trying to compare data obtained using incomparable methods.

I had initially planned to include visual stress and scotopic sensitivity in the thesis study and had asked participants to report visual stress and whether the font impacted this. I attended a conference on visual stress and came to realise that not only was it a topic that deserved a separate study but recognised that I did not have the necessary skills to draw valid and relevant conclusions. Thus all data pertaining to visual stress was discarded and the focus of the thesis tightened. Consequently, I deduced that constant and consistent reflection was vital to produce a coherent research study.

For me the most valuable aspect of any reflection is to make one's thoughts and writing clear, logical and comprehensible. By considering criticism and comments from knowledgeable others one's own skills can be developed. All research needs to be understandable to those who read it, especially, in this case, to those in education and allied professions. I have striven to make the report of my research not just clear in my mind but explained sufficiently clearly for everyone else to benefit from the findings.

Final stages

As I had been in poor health in 2018, which had threatened to derail my progress, I moved on quickly to present my thesis proposal and obtain ethics approval. I managed to recruit several schools willing to participate, with engaged staff and enthusiastic parents/carers and pupils. Luckily by the end of 2019 I had collected sufficient data for analysis and for my findings to show valid conclusions, as who could have predicted what 2020 would throw at us?

In March 2020 I was identified as clinically vulnerable and advised to isolate. This meant that commuting to UCL, IoE was no longer possible. All schools became a no-go area for me. I found this enforced isolation very difficult and struggled to keep progressing with my analysis and writing. In September 2020 my husband, who was a huge support to my studies, died suddenly, so I withdrew from the course for a year. Stepping back onto the programme was difficult as I was amazed at how much detail of my research I'd forgotten. Large swathes of notes could have been made by a complete stranger as I had no memory of them. Continuing without my greatest supporter was difficult. Every day I wrestled with whether to continue. But here we are, reflecting on unprecedented times for everyone and moving forward to honour all who struggled with the pandemic. We have all coped with obstacles, large or small, professional or personal. The pathway that I had envisioned when I began on the EdD programme vanished and who knows what the future will hold. I do know that I have learnt new skills and adopted a new clarity of thought and analysis. What I had imagined being a smooth motorway through the EdD experience became a rocky almost invisible uphill track but hopefully, the destination is in sight.

Chapter 1: Introduction

1.1 Rationale for study

"I heard that there is a font that makes children with dyslexia good readers."

This comment that was made to me by the parent of a pupil that I worked with in my role as a dyslexia support specialist formed the rationale of my Institution-Focused Study (IFS) (Broadbent, 2018). This statement had been made a few years earlier and this was the seed that sowed my interest in how font design could impact the reading performance of young readers with dyslexia.

Bessemans (2016) states that varying the font for young or struggling readers should be seen as an additional tool for supporting reading but is not the ultimate solution to reading problems. If the effect of a particular font were to prove beneficial then reading interventions could be supported and made more effective by the choice of a specific font. Thus, I wondered if I should be using a font that potentially was able to assist pupils with reading difficulties in my one-to-one support sessions as well as using a font designed to help readers with dyslexia when creating classroom resources. Should I also be recommending a specific font to parents and colleagues?

I could not think of any detrimental effect that a simple change of font could cause a young reader with dyslexia, but I did not want to recommend a font without clear supporting evidence. It appeared that there was very little published work that either supported the inclusion of a particular font design or refuted claims that specialist fonts were beneficial. Thus I resolved to research the subject so that I could decide whether to recommend font use based on quantitative and qualitative evidence rather than conjecture. Consequently, my ‘Postgraduate Diploma in Social Science Research Methods’ modules were based on devising a suitable research question and methodology to research font design and dyslexia. My Institution-Focused Study (IFS) developed this idea further by investigating the impact of two different fonts on the reading accuracy and comprehension scores of six Key Stage 2 (KS2) pupils with dyslexia and six typically developing KS2 readers as a control group. Encouraging results indicating increased reading accuracy for the group of participants with dyslexia obtained from these studies motivated my decision to design and carry out the current study.

An effective intervention for children with dyslexia is to read more (Felici, 2012) but a child who is struggling to learn to read may be reluctant to practice if this reinforces their conviction that they will fail. I surmised that if young readers could find a font that ameliorated some of the difficulties that they experienced when reading, then perhaps they would be motivated to practice reading more frequently and then might make improved progress. The slow and inaccurate reading often exhibited by my pupils contributed chiefly to slower than expected progress and they often became caught in the downward spiral of the Matthew Effect (Stanovich, 1986), whereby poor performance leads to an unwillingness to read, which in turn leads to falling further behind their classmates. If this cycle could be broken by a change in font that might help them to feel supported in their reading, then rather than lagging further behind, they might be able to progress and even begin to catch up with their typically developing peers. If a font such as OpenDyslexic enables a child to read more successfully then they might be induced to read more willingly, more regularly and also with more enjoyment. This could then lead to reading interventions being more effective through increased practice and possibly improved self-belief. It is unlikely that the font's effect would apply consistently to every reader, as the effects of dyslexia are not experienced uniformly (Kelly & Phillips, 2016), but it could be an extra line of support for those who experience a benefit. Consequently, the font could be instrumental in supporting the effectiveness of reading interventions and lead to greater reading progress and higher attainment scores.

1.2 Fonts designed for readers with dyslexia

Several fonts have been designed with the specific aim of ameliorating difficulties experienced by those with dyslexia. The primary intention of these specialist fonts is to improve legibility by having clear letter design, mainly with unique shapes, well-proportioned ascenders and descenders, clear negative space and clearly defined apertures. These include Easy Reading (Alfonsetti, 2013), OpenDyslexic (Gonzalez, 2012), Dyslexie (Boer, 2008), Sylexiad (Hillier, 2006), Lexie Readable (Bates, 2004), Read Regular (Frensch,2003). Lexie Readable and Sylexiad were designed primarily for adults with dyslexia, so were excluded from this current study with its focus exclusively on the reading performance of children. Read Regular was designed as part of a university thesis and then developed by Dutch publishing house Zwijsen, who now controls all rights to its use.

Easy Reading, Dyslexie and OpenDyslexic are readily available to download online. Easy Reading is used widely in Italy. It is free for personal use but requires a commercial licence for professional use. Dyslexie is free for home use but educational and business use are subject to a charge for a lifetime licence. It is downloadable on Windows and Apple but not on Chromebooks or mobile devices. OpenDyslexic is available to download with no mandatory fee but with an optional donation and has free updates. It is available for Windows and Mac plus iOS and Android devices. It is also available on e-readers such as Kindle and Kobo and included in an increasing number of apps. Due to its wider availability without additional cost implications, OpenDyslexic was chosen for this study so that if results indicated a benefit it could be recommended to the broadest selection of learners.

As it had already been shown in my IFS (Broadbent, 2018) that OpenDyslexic font had a beneficial effect on reading performance for pupils with dyslexia, I decided for continuity to use OpenDyslexic rather than change to or include another specialist font. Its use also allows assumptions and findings about the font outlined in the IFS to be developed by further research.

1.3 The design of OpenDyslexic

Gonzalez (2012) designed OpenDyslexic font with the goal of increasing readability for readers with dyslexia. The design aims to prevent the perception experienced by some with dyslexia that letters swap positions and words move. The unique letter shapes reduce the likelihood of letter confusion. In some fonts certain letters are mirror images or reversals of other letters, e.g when using Arial font ‘b’ and ‘d’; ‘p’ and ‘q’ are reflections and rotations of each other. In OpenDyslexic all letters have unique shapes so that no letter is the mirror image or reversal of any other to aid in avoiding misidentification e. g.

b d p q u n

The uniqueness of form is achieved by the use of thicker bases, slanted letters and varied length ascenders and descenders.

Figure 1.1. OpenDyslexic font

the quick brown fox jumps over the lazy dog
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG

The heavy weighted bases of each letter (figure 1.1) are intended to emphasise the correct orientation of each letter and to reinforce the line of text so that readers not only read along the correct line but can also locate the subsequent line more readily. The default spacing of OpenDyslexic is more expanded than many traditional fonts and this has the intention of preventing readers with dyslexia from experiencing the crowding effect, whereby the letters and words flanking the centre of focus interfere with correct perception and identification (Franzen et al., 2019). Theoretically, readers with dyslexia would be less likely to misread words by transposing peripheral letters from adjacent words into the word that they are reading. OpenDyslexic font was designed to enhance the reading experience of those with dyslexia. It is regarded as a compensating tool to support reading and not as a cure for dyslexia (Morley, 2018).

1.4 Anecdotal Evidence

Gonzalez (2012) asserts that the use of the font will mitigate some of the common reading errors caused by dyslexia. However, these claims are not backed up by data to substantiate how this statement has been arrived at. Information on the OpenDyslexic website (www.opendyslexic.org) gives links to some studies conducted but it is not possible to access the content of all of these studies via the links provided, so consequently it cannot be seen whether they provide supporting evidence for the font's efficacy.

An Internet search reveals magazine and newspaper articles that include quotes and anecdotal evidence concerning OpenDyslexic, e.g. Daily Mail, (8.10.2012); BBC online news (26.09.2012); Huffington Post online (10.07.2012). These articles coincide with the launch of the font so can also be regarded as publicity rather than any form of balanced or impartial recommendation. Burgess (2012) outlines comments from students aged 10 to 18 but approaches the subject from the perspective of journalism rather than research, as there is no information concerning sample recruitment and methodology. Broad answers such as "I can read this new font easier" and "I think I prefer the normal font" give balance but do not have the gravitas of peer-reviewed findings. Smith (2016) outlines that personal experiences based on what has been learned or understood cannot be used as evidence. Any comment indicating that a dyslexia-friendly font allows somebody to read "better" is not useful as it does not supply a measurable improvement. It is more closely aligned with personal inclination than confirmatory evidence. Recent praise of dyslexia-friendly fonts tends to be more measured and considered than in some earlier journalistic articles. This may be reflected in the target audience. Initially, this would have been the public at large but now reporting is more likely to be aimed at parents of children with dyslexia who tend to want practical support that gives a consistent benefit rather than sensationalism and possible false promises (Epstein, 2016). Those in education want to recommend proven support methods rather than fads (Saha, 2020). Eden (2017) and Morley (2018) argue that care should be taken before endorsing a product and postulate that reporting unsubstantiated testimonials for fonts may popularise the design before it has been studied in depth by controlled trials.

1.5 Scope and structure of study

This thesis builds on the findings of my IFS, which found that OpenDyslexic font had a significant effect on the reading accuracy scores of Key Stage 2 (KS2) pupils with dyslexia (Broadbent, 2018). The control group of typically progressing readers were not similarly impacted. Reading comprehension was not influenced by font design for either group. The participants varied in their preferred font. This thesis aims to investigate the impact of font in more depth and to include more participants so that the findings have greater validity and reliability. Additionally, as a potential measure of reading fluency, the reading rate of the participants was considered. The spacing that OpenDyslexic font has as a default setting was also controlled to ascertain whether it is the font design or spacing that is influencing any potential impact.

This study comprises a literature review in two parts: part 1, which considers learning to read and reading difficulties associated with dyslexia (Chapter 2), and part 2, which outlines current thinking on the impact of font design on reading and research on legibility of the printed word (Chapter 3). The methods (Chapter 4) of the current study are then explained, followed by results (Chapter 5) together with a discussion (Chapter 6), which highlights the key implications together with possible practical and professional recommendations.

Chapter 2:Literature Review (part 1):

Learning to Read and Reading Difficulties

The literature review for this study is presented in two parts: part 1, 'Learning to read and reading difficulties', which considers the processes involved when we read, how reading achievement can be assessed and the difficulties that are sometimes experienced when learning to read, i.e. dyslexia; and part 2, as Chapter 3 'Reading and typography', which includes font design, perception of the legibility of text, and the impact of font design on readers with dyslexia.

2.1 The brain and learning to read

Language is a method of human communication, either spoken, written or signed, whereby words are used in a structured and conventional way. The creation of the written word has made it possible to store large amounts of information outside the human mind and share it repeatedly across distance and time (Barclay, 2019). This necessitates the skill of reading, which is a multifaceted psychological process (Gentry & Ouellette, 2019).

Whilst not a specific focus of this study per se, it is interesting to consider how reading occurs in the brain, from perception of print to reading with meaning. The role of font and its legibility can be linked to how the brain learns to read by following the pathways from the perception of print by the eyes, which may be impacted by font design, through the processes and pathways necessary to link that visual input of graphemes to phonemes and also to access meaning.

Printed text must be perceived and processed visually; the perceived symbols matched to sounds of language and connected to meaning; speech accessed for pronunciation and higher order skills navigated to enable comprehension. This chain of events happens within milliseconds in proficient readers, who recognise the printed word automatically and seemingly effortlessly. Dyslexia, when reading does not happen so apparently effortlessly, refers to a learning difficulty that primarily affects the skills involved in decoding and spelling and is generally thought to include difficulties in phonological awareness, see section 2.6.

It is estimated that humans have been communicating by speech for over 100,000 years (Lilienfeld et al., 2018), which is a skill using innate instincts and intelligence and for most of human history existed without a written form (Seidenberg, 2017). Reading was invented at least five thousand years ago (Dehaene, 2009), so is a relatively recent cultural invention and consequently it is not regarded as a natural skill. The human brain has not evolved to read (Gabrieli et al., 2010) so the ability to read is not instinctive but needs to be taught (Coch, 2010). Multiple neural systems with their own specialism need to be accessed to actively construct a brain that can read (Dehaene & Cohen, 2007). The neurological reading circuit is spread out across the brain so an intricate series of connections need to be established for it to exist. Thus, successful reading relies on multiple processing areas, as shown in table 2.1.

Table 2.1 Brain Lobes and Functions (Gentry & Ouellette, 2019, p36)

Lobe	Noted Functions	Contribution to reading
Occipital	Visual processing	Processing letter shape; visual scanning
Parietal	Sensory processing; bodily awareness; understanding	Processing sound sequences; linkage to syntax and meaning
Temporal	Auditory processing; language; emotion	Phonemic processing; linkage to syntax and meaning
Frontal	Motor output; speech and language; thought; self-regulation	Speech planning and grammar/morphological processing

2.2 Visual processing and reading

Reading is principally a linguistic experience (Seidenberg, 2017) but the visual symbols on the page have to be perceived by the eyes before the text can be processed to obtain meaning, so as Dehaene (2009) and Seidenberg (2017) assert, reading starts with the eyes.

2.2a Saccades and fixations

Reading starts with visual processing. The centre of the retina, the fovea, has a dense concentration of high-resolution receptor cells fine enough to recognise print. Visual accuracy is optimal at the centre and decreases towards the periphery (Dehaene, 2009). The retinal periphery recognises general shapes and outlines and is good at localizing objects and detecting movement and flicker (Evans, 2001). So that text can be scanned with the most sensitive part of our vision, i.e. the fovea, the eyes need to be in motion (Dehaene, 2009). Our eyes do not travel continuously across the page but in small steps or saccades (Rayner, 1998). We do not perceive any discontinuities and the text appears seamless as we move our eyes from one location to the next (Taylor & Perfetti, 2016).

After each saccade the eye fixes on one spot to read, a fixation, and while the fovea is analysing the small group of centrally fixed letters, the peripheral retina is processing the general shape of letters and words ahead to determine where the next fixation will be (Evans, 2001). The duration of a fixation depends on the difficulty of a word, the difficulty of the text, the attention that the text needs to be given and the skill of the reader. For a skilled reader, the average fixation lasts 200 milliseconds (Sousa, 2014), with children's fixations tending to be for longer (Joseph et al., 2009).

Information is only taken in at each fixation. The number of letters seen in the region of effective vision or perceptual span is limited. The perceptual span for skilled readers consists of 3-4 letters to the left of a fixation and 14-15 letters to the right in left to right reading systems (Bélanger & Rayner, 2015). The perceptual span is asymmetrical to the left if the language read is scanned from right to left e.g. Hebrew, and saccades are shorter where character density is greater e.g. Chinese logograms. Thus readers adapt scanning text to their own language and script (Rayner et al., 2010). Beginning readers and readers with dyslexia have smaller perception spans extending approximately 5 to 11 letters to the right of the fixation when reading in English (Häikiö et al., 2009). The amount that can be seen in one fixation seems also to be linked to how rapidly sense can be made of it (Seidenberg, 2017). Thus, limitations to the width of the perceptual span are cognitive as well as physical as text needs to be perceived, decoded and comprehended. A small span width on a densely packed script will require the equivalent cognitive effort as that needed to understand a wider span on a language system that uses a less densely packed script. Consequently, the size of saccades is not just driven by visual perception but also by how much information can be processed at any one time.

Not all saccades progress along the line of text. Sometimes the forward eye movement is interrupted by a regressive saccade whereby the eyes return to a previous word or location to reprocess text already perceived (Vitu, 2005). This might occur if text has been misread; to enhance understanding; correct misunderstanding; or clarify ambiguity (Seidenberg, 2017). Developing readers make regressive saccades at a higher rate than skilled readers, often to correct decoding errors or failures in comprehension (Joseph et al., 2009).

Readers with traits of dyslexia tend to make excessive saccades and fixations (Franzen et al., 2021). It is suggested by Benfatto et al. (2016) that it is possible to identify young readers that may be in danger of persistent literacy difficulties by using eye tracking to identify atypical eye movements when reading, which they assert relates to the level of facility that young readers may have to process grapheme-phoneme correspondences. De Luca et al. (2002) report in their study of 12 participants with dyslexia and 10 age-matched controls that typically developing readers tended to adjust forward saccades to word length, which had the advantage of allowing whole words to be processed. Conversely, participants with dyslexia increased the number of saccades when processing longer words with the saccade amplitude remaining small, which implied that word sub-units were being analysed with words of around eight letters being read in 3 to 4 segments. This indicated that those with traits of dyslexia were using a sub-lexical grapheme to phoneme procedure to decode the target words.

Readers with dyslexia also tend to make more regressive saccades than typically developing readers. Wong (2020) suggests that this could be due to the fact that if such readers read inaccurately then they need to re-read material more often than good readers. She argues that making more fixations may become a habit, as the number of regressions tends to persist even when the text comprises material that is relatively easy to decode and process.

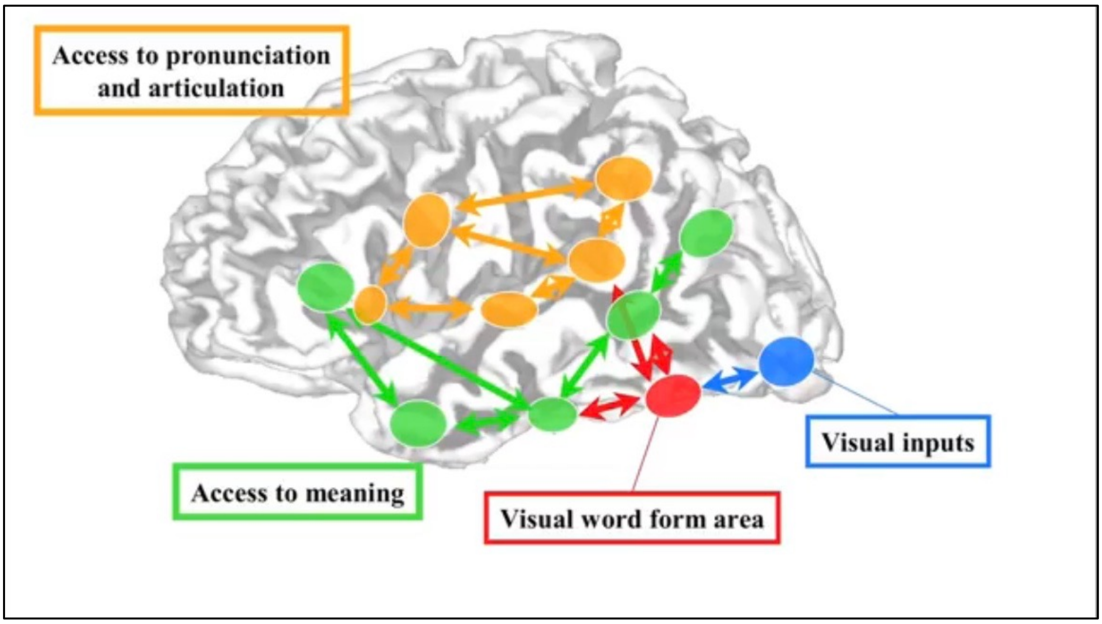
When the end of the line of text is reached a return sweep is necessary to move the attention to the start of the next line. This return sweep may undershoot the target, necessitating a corrective saccade towards the left margin (Parker et al., 2020). Early readers and readers with dyslexia can be easily disrupted at reading line breaks. Typically developing readers tend to overcome this during Key Stage 2 but return sweeps locating to the wrong line or incorrect location on the consecutive line often persist in readers with dyslexia (Tiffin-Richards & Schroeder, 2018).

2.2b. The neurological reading circuit

Once the eyes have received the visual information this input travels to the visual cortex in the occipital lobe for processing (Figure. 2.1). From here, according to the two streams hypothesis (Goodale & Milner, 1992), the cortical visual processing stream divides into two. Information concerning location and spatial properties passes along the dorsal pathway to the parietal lobe. Information concerning the identification of what is perceived moves along the ventral pathway to the temporal lobe, processing detail of colour, texture, pattern and fine detail. This enables features of letters to be processed.

Once letters are identified the information moves to the left fusiform gyrus area, known as the Visual Word Form Area (VWFA) (Figure 2.1). Information that is identified as shapes and forms other than letters, is processed on the right side of the brain in the Fusiform Face Area (FFA).

Figure 2.1 Cortical networks used for reading. (Dehaene, 2009, p63)



It is hypothesised by Dehaene and Cohen (2011) that the VWFA is highly specialised in recognising certain types of shapes that occur naturally in the environment but which are now associated with written language. Before learning to read the VWFA area reacts to faces and objects but with the onset and development of literacy the response towards recognising letters and letter strings becomes increasingly dominant and the response to faces and objects decreases. The recognition of faces and forms eventually moves predominantly to the FFA.

The original function of recognising faces and objects allows them to be identified regardless of their left or right orientation. This is an evolved, pre-existing ability of the cortical region (Dehaene, 2014). Once we know the visual shape of an object we immediately extend that knowledge to its mirror image, known as mirror invariance. This ability needs to be modified when learning to read so that the directional aspects of letters are distinguished. This perception process of suppressing mirror invariance and distinguishing the orientation of similar forms must be achieved when learning to read. Letters that differ only in orientation e.g. ‘b’ and ‘d’ need to be seen as a different rather than as the same object facing in an alternate direction. The suppression of mirror generalisation becomes an active process in literate individuals (Dehaene et al., 2015). Difficulties in detecting orientation may linger in some young readers longer than average, accounting for some children confusing direction in letters and reading or producing letter reversals (Ahr et al., 2016). Dehaene (2009) asserts that this anomaly shows that the brain was never intended for reading but uses and alters pathways that had evolved for a different purpose. Before formal schooling a child’s brain has a symmetry constraint that exerts a strong influence until the visual brain determines that unlike with most objects, it is vital that left and right orientations are perceived when looking at letters.

Learning to read requires the construction of an efficient connection between visual areas and language areas. The VWFA acts as an interface connecting vision to spoken language and meaning. This is shown schematically in Figure 2.1, where the blue arrow indicates how vision is linked to the VWFA and then, via the pathways shown in red, to the regions of the brain shown in yellow that are concerned with speech and the pronunciation of spoken language and the regions in green which indicate the areas in the brain used in retrieving meaning of language. The pathways in the brain depicted in green and yellow in Figure 2.1 are not specific to reading but contribute primarily to spoken language. The connections are bidirectional and more intricate than indicated on the diagram.

Neuroimaging studies have shown changes in brain activity as children learn to read. Maurer et al. (2006) reported no specialised activation relating to sensitivity to letter strings in pre-readers but increased activation within the VWFA within two years of literacy instruction. This assertion was substantiated by Brem et al. (2010) who identified increased sensitivity to print with activation of the VWFA following targeted interventions via software that trained grapheme-phoneme correspondences with 32 kindergarten pupils prior to learning full word reading. This demonstrates that the plasticity of the brain responds to reading instruction by the gradual recruitment and increasing involvement of the VWFA. Monzalvo et al. (2012) compared fMRI activations to written words in 23 children with dyslexia and 23 with typical reading development, aged 8 to 10 and found that there was disorganisation in the ventral vision and spoken language areas of the children with dyslexia. The participants with dyslexia showed expected responses to pictures of houses and checkerboard patterns but had a reduced activation to words in the VWFA and faces in the FFA, indicating that the reaction to faces had not yet transferred from the left fusiform gyrus to the right to the extent expected.

2.3 Routes to reading

Before children learn to read they acquire vocabulary by listening to others, then practising pronunciation when they learn to vocalise speech and practice new words in conversation. These learned, familiar words become part of the orthographic lexicon (Martinet et al. 2004). Once literate, the lexicon can be subdivided into categories defined as 1) orthographic lexicon, which stores knowledge of visual forms i.e. the spelling of familiar words, recognition of letters, graphemes and morphemes; 2) the phonological lexicon, which stores how words are pronounced; 3) the semantic lexicon which stores meanings of words (Coltheart, 2006).

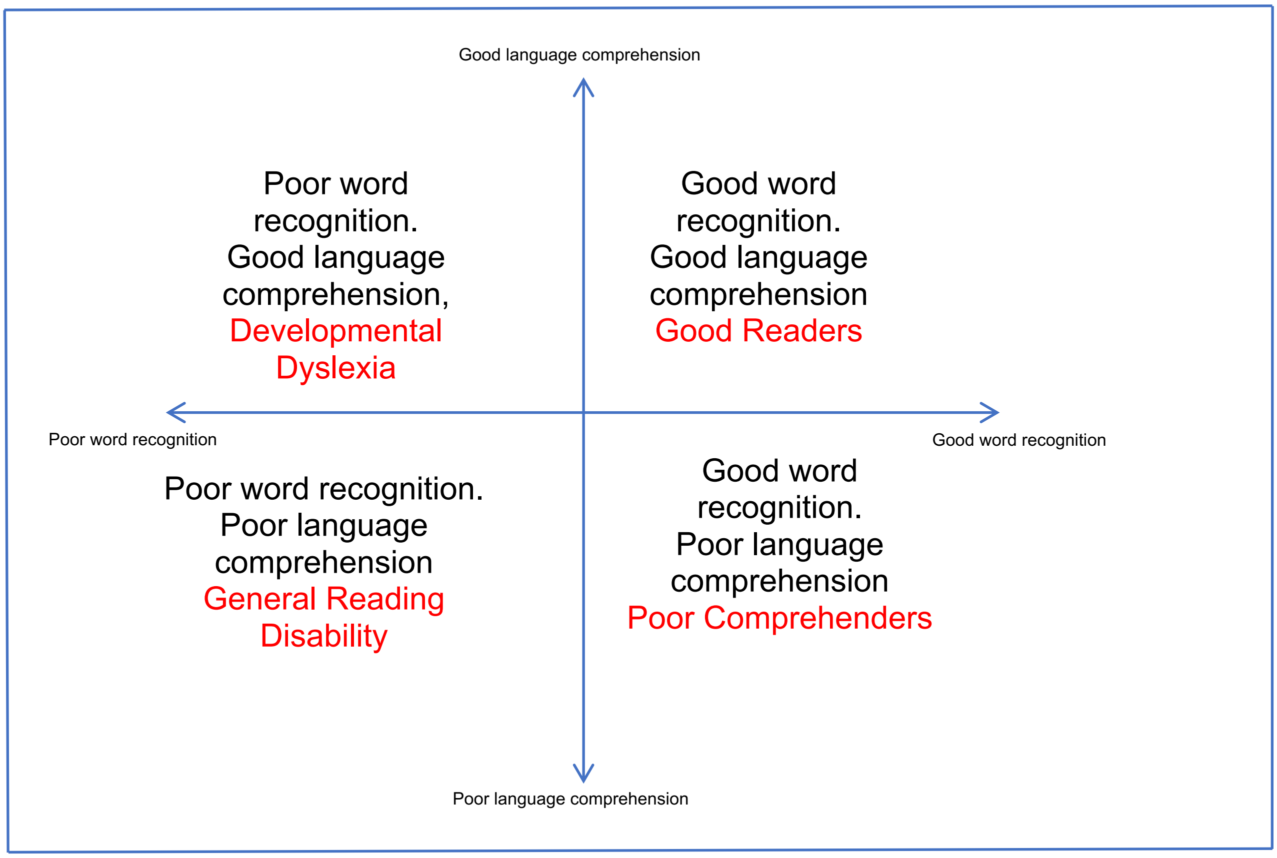
Reading at word level is theorised as depending on two distinct processes, a lexical route, in which words are recognised as a whole, and a sub-lexical route, which is focused on units smaller than words (Coltheart et al., 2001). If the reader is skilled and the word familiar, then phonemes do not need to be sounded out so the VWFA processes the word via the lexical route that identifies the letters and indicates that it is already in the orthographic lexicon. Meaning is assigned directly to the word form from the semantic lexicon and then the phonological lexicon retrieves pronunciation. If the word is unknown then the sub-lexical route is followed whereby the pronunciation is assembled using grapheme to phoneme mapping and if the word is then recognised meaning is accessed from the semantic store (Sousa, 2014). Much early, developmental reading follows the sub-lexical route but as readers become more skilled then increasingly more text is accessed via the lexical route. The two routes work in close co-ordination as if a word is novel or rare it is still accessed by the sub-lexical route even in skilled adult readers.

As can be seen from the sub-lexical reading route, a key process in learning to read is phonological awareness, the understanding of how the spoken word is comprised of smaller sound units, e.g. the word dog broken into its component sounds /d/o/g/. These sounds must then be mapped onto the orthographic symbols or graphemes (Coch, 2010). Phonemes can then be manipulated to form a new word, e.g. if the initial phoneme /d/ from ‘dog’ is replaced by /f/ the resulting word is /f/o/g/ ‘fog’. Once a child can identify phonemes and reassemble, reorder and manipulate them then they are said to have achieved phonological awareness. Phonological awareness is positively correlated with reading skills and a strong indicator of future reading success (Snowling, 2000).

2.4 The ‘Simple View of Reading’

Reading is not merely decoding phonemes to create words or accessing words from the orthographic lexicon to construct sentences. The fundamental function of reading is to extract meaning, i.e. to comprehend (Cain, 2010). The relationship between deciphering text and attributing meaning is theorised by several models, but a detailed analysis of these is beyond the scope of this study. Whilst not without its critics, the Simple View of Reading (Gough & Tunmer, 1986) is often presented in government documents and school policies to explain the interplay between decoding and reading comprehension. Thus this theory is well known in the school environment and is added here as background information to indicate why measures of decoding and comprehension were included in the research methodology.

Figure 2.2 The Simple View of Reading, adapted from Cain (2010) and Rose (2009)



The model attempts to simplify the complex process of combining the cognitive skills of decoding with the ability to comprehend a text (Kirby & Savage, 2008). The Simple View of Reading states that reading is a product of decoding and comprehension; R = DxC, where R equals reading comprehension, D equals decoding and word recognition and C equals language comprehension. Therefore, if either D or C is zero, i.e. decoding or comprehension is not taking place, then R must equal zero, i.e. reading is not happening. If the child cannot decode, or, if they can decode but not comprehend, then, by definition, they cannot read. Thus, both decoding and comprehension are interdependent and essential to reading.

Figure 2.2 shows how word recognition and language comprehension can be measured along a continuum and intersect to predict varying groups of readers. Readers with dyslexia are considered to have good language comprehension but poor word recognition, whereas typically progressing readers exhibit both good language comprehension and good word recognition skills.

The use of the term decoding may be ambiguous. It could be used to mean the ability to sound out words using phonic rules, or it can mean successful word recognition influenced by phonology, orthography and semantics by whichever route (Plaut, 2005). In this study, it will generally be taken to mean the latter.

2.5 Reading assessment

Achieving good reading skills is a fundamental goal for all children, as it is the key to learning (Muter, 2021). Assessment of the progress of young readers on their journey to reading competence is usually carried out by monitoring reading accuracy and reading comprehension. Additionally, reading rate is often considered a gauge of reading fluency and an indication of whether decoding is becoming automatic. These aspects of reading assessment are now considered.

2.5a Reading accuracy

To achieve reading accuracy, i.e. to be able to read without errors, the reader needs to have phonemic awareness and knowledge of phonics (Blevins, 2017). This enables the decoding of words, which allows the development of word recognition. The initial phase of learning to read based on phonology and decoding tends to be slow and laborious with errors arising from mismatching graphemes and phonemes. The consistency of the mapping of phoneme to grapheme varies between languages. English is a language with an opaque or deep orthography, where letters are not consistently matched to sounds or where letter combinations can have more than one pronunciation. Languages with a transparent or shallow orthography e.g. Finnish, Italian, or Dutch, have more regular grapheme-to-sound correspondences (Seymour, 2005). To read accurately a reader must be able to assign the correct sound to each letter or letter group and build words from phonemes and also maintain learned words in the mental lexicon. Assessment can be by a single-word reading test, but as reading is mostly in the context of continuous text, accuracy should also be assessed using meaningful prose (Clay,1991).

Accuracy of decoding is often used within schools to move children through reading schemes. Clay (1991) argues that a passage or book has been mastered once reading accuracy reaches 95%. This indicates that the reader has consolidated the skills required to decode the text and has the ability to achieve the new skills required for the next level or book in a reading scheme. A small percentage of inaccuracy should be allowed before moving to the next level of difficulty so the reader continues to be challenged by the gradient of text difficulty and does not face a step from an easy text, 95+% accuracy level, to a hard text read with less than 89% accuracy (Clay, 2000). A text with some difficulties allows the reader to realise that encountering challenging text develops reading skills (Hodgkinson & Small, 2018). Rodgers et al. (2018) argue that any texts read with less than 90% accuracy should be avoided as damaging to reading progress, with the negative impact increasing the more instances of reading at less than 90% accuracy that the student experiences. This is especially important for struggling and beginner readers who need to experience high levels of reading accuracy, using texts that are relatively easy for them, so that decoding skills are supported and fluency encouraged (Shanahan et al., 2012).

Reading accuracy is usually assessed by either phonics checks of real and nonsense words or by the child reading coherent, meaningful text. The oral reading performance is analysed by recording and reviewing errors on running records (Clay, 2000) or by interpreting miscues (Davenport, 2002; Goodman, 1969). These methods are often regarded as interchangeable, providing similar data that allows teachers to assess progress and allocate suitable reading materials. However, Harmey and Kabuto (2018) argue that the two approaches of running records and miscue analysis quantify different aspects of the reading process and are underpinned by different theoretical concepts. All uncorrected errors recorded on a running record are considered errors and determine reading accuracy, whereas miscues can be analysed further to show whether an incorrectly produced word is grammatically and semantically acceptable and thus indicates a level of comprehension as well as accuracy. It could be argued that to enable pupil progress to be tracked consistently the same tool should be used, as different assessment methods, which appear to be similar, can produce different or conflicting results. Conversely, the use of a variety of assessment tools could give a more complete picture of the complex process that has to take place for a child to learn to read.

Types of reading error identified when assessing oral reading performance can provide an in-depth picture of a child's reading development. Davenport (2002) asserts that the most common type of reading error is the substitution error whereby a word is substituted by a word that exists in the language not a meaningless non-word. The sense of the sentence can be preserved or could be altered depending on the word substituted and the context of the substitution. Clay (2001) indicates that when a word is read incorrectly the wrong word is influenced by syntax, meaning, letter knowledge and letter-sound relationships. A child may cue only the first letter of the word and then use that sound to guess and make a substitution error. e.g. 'on' in place of 'out'[2]. Errors may preserve the first and last letters e.g. 'your' – 'year' in what can be called a scaffolding error (Ehri, 2005). Savage et al. (2001) assert that scaffolding errors that preserve the beginning and end letters correlate to later word reading success. Stuart and Coltheart (1988) assessed errors and found that with increased phonemic awareness and letter knowledge, words read in error tend to share letters with the target word, whereas errors that bore less resemblance to the target word indicated weaker phonological skills. An example of this is the transposition of 'forming' for 'foaming', which has only one letter incorrect, indicating that the grapheme-phoneme correspondences are almost established, whereas 'rushed' substituted for 'reddish' implies poor grapheme-phoneme correspondences. Some readers may transpose a letter group from the end of a word to the beginning

[2] All examples obtained when testing with YARC texts during this study

e.g. ‘very’ in place of ‘never’. This indicates that the reader still has to develop full left to right parsing of words (Savage et al., 2001). Non-phonologically linked errors do not correlate positively with reading ability as they indicate that the reader is not making grapheme-phoneme links as a decoding strategy. Errors may instead be connected by context e.g. ‘trunk’ in place of ‘stump’ when reading about a tree.

Additional error types that may be recorded in reading accuracy assessments include additions, where an extra word is inserted; omissions, where a word is left out; reversals, where the position of letters or words are reversed; mispronunciations, where a word is wrongly pronounced, or partially or incorrectly decoded to create a word that has no meaning; refusal, where the pupil is unable to attempt the word (Snowling et al., 2009b). The number and type of reading errors can be analysed and interventions devised to give targeted supplementary assistance for readers in danger of falling behind expected levels (Phillips & Kelly, 2018). To assess reading accuracy, techniques such as miscue analysis (Goodman, 1969) and running records (Clay, 1993, 2000) can be beneficial.

2.5b Reading fluency

Poor reading fluency is a more salient difficulty in transparent orthographies where phoneme-grapheme correspondences are regular (Elliot, 2020). Nevertheless in an opaque orthography such as English, in addition to being able to decode words accurately or identify and access words already in the mental lexicon, a reader should also develop reading fluency. Although fluency is usually assessed by measuring reading speed, Rasinski and Cheesman Smith (2018) state that approaches to fluency that stress reading at a prescribed reading speed do not encompass all that is encapsulated within fluency. They assert that fluency should demonstrate automatic word recognition, rhythm, expression and smoothness. Wolf and Katzir-Cohen (2001) outline many definitions of reading fluency but conclude that fluent reading should be smooth, accurate and automatic.

Automaticity in reading is outlined by Kuhn et al. (2010, p.231) as exhibiting “speed, effortlessness, and a lack of conscious awareness”. It relies on visual and auditory processing, so for automaticity to occur these processes must become synchronized (Breznitz, 2005). In the early stages of learning to read, readers may be accurate but slow and inefficient at recognizing words (Armbruster et al., 2001). Gradually word recognition becomes more automatic and should show a speed similar to speech (Quigley, 2020). The pace that the reader uses gives an idea of the effort required to decode, with children who struggle with decoding and who lack fluency, reading slowly and laboriously (Sousa, 2014). A text will often not be read at one consistent pace, as even a skilled reader will slow down when encountering unfamiliar vocabulary or topics (Breznitz, 2005). Sousa (2014) argues that automaticity and fluency are not the same, and asserts that automaticity is essential but not sufficient for fluency, with a further component, namely prosody, required to produce fluency.

Prosody is defined as expressive reading with appropriate rhythm and cadence, plus phrasing that reflects the meaning of the text (Rasinski & Cheesman Smith, 2018). A role that prosody plays in assisting with understanding of text is to help the reader by chunking text into syntactically appropriate and meaningful phrasal groupings of words. Not all phrase boundaries are marked by punctuation, so readers must use prosody to parse text into appropriate phrases. Alvarez-Cañizo et al. (2015) investigated reading fluency and comprehension with a sample of 40 children from grade 3 (age 8) or grade 6 (age 12). Half of each group was assessed to have good comprehension both orally and of written texts, and half had good oral comprehension but poor comprehension of written texts. Results showed that the group with lower levels of reading comprehension made more inappropriate, ungrammatical pauses, so it was concluded that some reading comprehension problems were related to a lack of development of good reading fluency, indicated by poor prosody. Thus, it can be argued that by improving reading fluency reading comprehension can be enhanced, especially when oral comprehension is already well-developed.

Conversely, Applegate et al. (2009) assert that fluent reading does not always lead to improved reading comprehension. Participants, who were judged to be strong readers because of decoding accuracy and high fluency test scores, were assessed for recall and high-level comprehension. Over 30% of those tested achieved relatively low reading comprehension scores, below the expected mean for their age group despite being identified as fluent readers, thus indicating that a high level of fluency had not produced good reading comprehension. Thus, it remains unresolved as to whether reading fluency is a facilitator of reading comprehension.

2.5c Reading comprehension

The fundamental reason for reading is to extract meaning. The Construction-Integration Model (Kintsch & Rawson, 2005) can be used to explain this process by proposing three levels of skills required to comprehend text. The first of these is the linguistic level whereby the reader recognises and processes individual words and accesses their meaning. Thus, the reader not only needs a vocabulary of words and their meaning but also needs to be able to retrieve words from their mental lexicon (Perfetti, 2007). Developing an extensive vocabulary enables readers to understand more words, which supports comprehension. It is also important for young readers to understand how seemingly simple words can act as cohesive ties, i.e. how words link sentences and ideas together (Quigley, 2020). Anaphors, words such as “this”, “she”, or “it” may seem simple and accessible and a skilled reader can understand how sentences are connected by these words, but poor readers with less grammatical knowledge and practice may miss their meaning and function and interpret text incorrectly. Connectives such as “and”, “but”, “so” signal how to integrate the meaning of clauses. To make these connections the reader must hold the word that has just been read in working memory and consecutively make cohesive links with other words within the sentence or passage (Gathercole & Baddeley, 1993).

This leads to the second level in the model, the microstructure, which is when the reader can assimilate the meaning of individual words to form meaning at the sentence level or, by establishing that the meaning of adjacent sentences is linked, at the paragraph level.

In the third level, the macrostructure, the reader can incorporate these sentences to access meaning and identify themes, topics and genre information from the text (Kintsch & Rawson, 2005). These three levels form the textbase, which needs to be coherent and accurate so any inconsistences can be noted and resolved. Yuill and Oakhill (1991) posit that young children often judge texts to be fully comprehensible even if there are gross inconsistencies. The ability to fully monitor comprehension usually develops with age and reading experience, which indicates that creating a coherent textbase is a skill that takes time to develop.

If reading is unproblematic then readers remember the gist of the text, which is the main ideas, themes or plot. Precise wording is rarely remembered (Caccmise et al., 2008). The structure created in the mind to secure an understanding of text is named differently according to two theories, a 'mental model' (Johnson-Laird, 1983) or a 'situation model' (Kintsch, 1988). This study is not dependent on the precise difference between the theoretical stance of the two theories so "the construction of a coherent and integrated representation of meaning" (Oakhill & Cain, 2018 p.684) will be referred to as a situation model. This is described by Willingham (2017 p.119) as "representing the big picture, a memory constructed to encompass the overall situation described in the text". It is also influenced by information outside the text, namely other relevant knowledge that is remembered from prior experience and retrieved from long-term memory.

Prior knowledge can have a large influence on understanding. Sadoski (2008, p. 40) demonstrates this using the sentence "The guard invaded the paint for a dunk". It is possible to understand every word in that sentence but still be unable to comprehend it. However, a basketball fan, because of their prior knowledge of vocabulary and terms associated with the sport, would understand it immediately as meaning “the basketball player (the guard) ran into the area under the basket bounded by the foul lanes and the free throw line (invaded the paint) and leapt up and delivered the ball directly into the basket (for a dunk)” (Sadoski, 2008, p.41). Recht and Leslie (1988) studied the effect of prior knowledge on reading comprehension and found that knowledge of a topic had a much bigger impact on comprehension than general reading ability. This can link additionally to the issue of required knowledge outside a child’s cultural experience. This is especially important to learners with English as an Additional Language, who as well as building up a lexicon in English need to develop cross-cultural background knowledge (Khataee & Davoudi, 2018).

In addition to accessing prior knowledge the ability to generate inferences is important in fully understanding text. Text rarely includes every detail, so the reader is expected to use pieces of information explicit in the text to arrive at further information which is implicit and not spelt out. The various categories of inference type can be divided loosely into inferences that are generated from the text, and those that require additional information from the readers’ own understanding and real-world knowledge (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 1999; Cromley & Azevedo, 2007; Graesser et al., 1994; Kispal, 2008; Pressley & Afferbach, 1995; Snowling et al., 2009a). Inference made from the text includes making coherent links between words at a local level such as pronoun resolution and resolving incongruities or ambiguities within and between sentences. Inference requiring additional information includes applying real-world knowledge to go beyond what is stated explicitly in the text. The ability to deduce what is expressed implicitly is often needed to acquire a full understanding of the text. Young readers are capable of making inferences but are less likely to do so than older readers (Oakhill et al., 2015). This can result in developing readers often failing to realise that inference is needed to maintain global coherence. Kendeou et al. (2008) argue that children are more likely to make inferences when asked to elaborate on the text than when they have to establish coherence for themselves. This indicates that inferences that should have been made whilst reading the text may actually only be made when questioned afterwards.

It is possible that not all comprehension tests tap into the same set of skills to the same degree. This means that results may not be comparable. Bowyer-Crane and Snowling (2005) found that less skilled comprehenders performed better using WORD (Wechsler,1990) than NARA II (Neale,1989). Their scores when tested using WORD tended to fall within the average age range unlike scores obtained from NARA II. A large proportion of the questions on WORD require an understanding of literal information whereas NARA II questions tap into cohesive and knowledge-based inferences. Children who have difficulties in reading comprehension are generally most likely to struggle with inference so this can be masked by an emphasis on questions that can be answered by accessing literal information and recall and could explain higher scores when tested using WORD. Tests that ask more inference-based questions that require real-world knowledge to answer appropriately are more difficult for poorer comprehenders, so may penalise them and cause them to achieve lower scores on tests such as NARA II.

The difference in scores generated by reading comprehension tests was also researched by Colenbrander et al. (2016), who considered whether the reading comprehension test NARA (Neale, 1999) and YARC (Snowling et al., 2009b) returned comparable test scores. Results indicated that more participants were diagnosed as poor comprehenders when tested using the NARA than the YARC. Unlike the Bowyer-Crane and Snowling (2011) study, Colenbrander et al. (2016) did not consider the type of comprehension questions that were asked, so were not able to say whether or not the difficulties lay with inference-type questions. The results emphasise that reading comprehension results can vary between tests.

It can be seen that the assessment test results of reading accuracy, reading fluency and reading comprehension can be used to give a clear, detailed measure of a young reader's progress. To ascertain whether a change of font design might impact this progress, it is necessary to use a research tool that addresses these three elements. A clear overall picture would be achieved by employing a valid measure of assessing the entire spectrum of reading performance, namely by investigating any effect of font on reading accuracy, reading rate and reading comprehension.

2.6 Definitions of Dyslexia

Not all readers develop their reading abilities as expected. The literacy difficulties of these struggling readers are characterised typically by poor decoding skills resulting in inaccurate reading, which may lead to poor comprehension (Phillips & Kelly, 2018). These children are often regarded as having Dyslexia or exhibiting traits of Dyslexia.

Dyslexia is predominantly acknowledged as being a specific learning disability that is neurological in origin (Brady, 2019). It is characterized by difficulties with accurate and fluent word recognition, poor decoding abilities and poor spelling (Lyon et al., 2003). There is no universal definition of dyslexia that provides a definitive statement of what is meant when the term dyslexia is used. There are commonalities in published definitions but also an emphasis may be placed on different characteristics. Many organisations have their own definition of dyslexia, and this is especially so with national associations and dyslexia support charities. Organisations in several English-speaking countries including USA, Australia and Canada use the definition outlined by the International Dyslexia Association (IDA) as a basis for defining what is meant when the term dyslexia is used officially. The definition used British Dyslexia Association (BDA), from the work of Rose (2009), underpins this study. These definitions, together with the definitions used by Dyslexia Scotland, the Dyslexia Association of Ireland and that outlined by the Ministry of Education, New Zealand are compared in Table 2.2.

Table 2.2 Comparison of definitions of the term dyslexia used by different organisations

	British Dyslexia Association	Dyslexia Scotland	Dyslexia Association of Ireland	International Dyslexia Association	NZ Ministry of Education	Stichting Dyslexie Nederland
Reading difficulties	Dyslexia is a learning difficulty that primarily affects the skills involved inaccurate and fluent word reading and spelling.	…difficulties learning to read, write and spell	…learning difficulty affecting the acquisition of fluent and accurate reading and spelling skills.	It is characterized by difficulties with accurate and /or fluent word recognition and by poor spelling and decoding abilities.	…evident when accurate and/or fluent reading and writing skills … develop incompletely or with great difficulty…	… problems in learning to read accurately and fluently and/ or problems learning to spell…
Phonological deficit	Characteristic features of dyslexia are difficulties in phonological awareness…	…associated difficulties such as…phonological awareness…	…characterized by cognitive difficulties in phonological processing…	These difficulties typically result from a deficit in the phonological component of language…	…particularly phonological awareness…	Students have difficulty with sound processing and swap sounds and letters
Continuum of difficulties	It is best thought of as a continuum, not a distinct category and there are no clear cut off points.	…varies in degree…	Dyslexic difficulties occur on a continuum from mild to severe…		a spectrum of specific learning difficulties…	
Persistent difficulty	…persistence of dyslexic difficulties…	It is… life-long…	…occur across the lifespan and may manifest themselves in different ways at different ages.		...these difficulties are persistent difficulties…	… dyslexia is for life…
All intellectual abilities	Dyslexia occurs across the range of intellectual abilities.	Dyslexia exists … across the range of abilities…			People with dyslexia are found across the achievement spectrum…	
Discrepancy in performance		These difficulties often do not reflect an individual's cognitive abilities and may not be typical of performance in other areas.	…being unexplained in relation to an individual's other abilities…	…often unexpected in relation to other cognitive abilities...	[opportunities]… that are effective and appropriate for most other children.	Reading scores lag behind other school subjects, compared to peers and what can be expected in view of the child's intelligence (see below)
Educational experiences		…according to learning and teaching environment…	…despite access to appropriate learning experiences…	…the provision of effective classroom instruction…	…despite access to learning opportunities…	
Additional difficulties	Co-occurring difficulties…memory and verbal processing speed, …motor co-ordination, mental calculation, concentration and personal organisation… Visual and auditory processing difficulties	Associated difficulties…auditory and/or visual processing, oral language skills, short term and working memory, sequencing and directionality, number skills, organizational ability	Including difficulties in working memory, rapid naming, automaticity of basic skills…difficulties in organisation, sequencing and motor skills	Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge.	Difficulties with auditory and/or visual perception, planning and organizing, short-term memory, motor skills or social interaction.	
Response to intervention	The extent to which learners respond to well-founded intervention indicates severity of difficulties	Learners with dyslexia will benefit from…appropriate intervention and targeted effective teaching	…may be alleviated by appropriate intervention		…early intervention, highly individualised, skilled teaching focused on written language… is critical	Reading and spelling problems do not disappear even after intensive intervention.

Dyslexia is found in all countries, cultures and languages. The orthography of the language spoken and read is thought to determine how dyslexia manifests. In a language with a shallow orthography e.g. Finnish or Dutch, where there are fewer variations in letter/grapheme to sound/phoneme correspondences, dyslexia often exhibits as slower, less fluent than expected reading. In a more opaque orthography, such as English or French, in which a letter or combination of letters may have more than one pronunciation, dyslexia becomes apparent through inaccurate decoding. The definitions initially considered were Anglo-centric so to compare a different language with a different orthography the definition used by the Stichting Dyslexie Nederland[3] was included.

In Table 2.2 all definitions include the fundamental premise that dyslexia is a reading difficulty that has a basis of phonological difficulties. Other elements vary between definitions. In the past the term “dyslexia” was used to distinguish poor readers with higher intelligence than was reflected in reading ability, as indicated by reading age and IQ test scores. Poor reading would contrast unexpectedly with other skills and the label dyslexia would be applied. Poor readers, who achieved lower IQ scores, were regarded as not having dyslexia as their poor reading was “expected” alongside poor ability in other areas. However, this IQ discrepancy model has been discounted (Siegel & Hurford, 2019), as it has been postulated that the same core difficulties exist regardless of whether there is a discrepancy between cognitive ability and reading achievement. The ability to decode the written word is independent of general intelligence (Joshi & Aaron, 2008). By disregarding this discrepancy, all children who are identified as struggling readers will be more likely

[3] I speak fluent Dutch so read and translated the original to include on table 2.2

to be provided with remedial interventions than just those with a label of dyslexia based on the identification of an IQ discrepancy. There are arguments put forward for dyslexia to be synonymous with poor reading so that all children who struggle to learn to read are eligible for funding and intervention (Cooke, 2001; Elliot, 2020; Elliot & Grigorenko, 2014). The Stichting Dyslexie Nederland now includes an addendum to their definition to emphasize that the idea of there being a discrepancy between reading level and general intelligence has been discounted and is no longer a factor in identifying dyslexia.

The DSM-5 diagnostic criteria now eliminate the IQ achievement discrepancy and replace it with four criteria, all of which must be met before a specific learning difficulty can be diagnosed. These criteria include identifying reading difficulties that have persisted for at least six months despite targeted intervention and which cause academic impairment (Tannock, 2014). The IDA, Dyslexia Scotland, and DAI mention discrepancies which refer to discrepancies in educational performance that are due to reasoning and processing differences, i.e. those with dyslexia may be capable of reasoning and understanding but may have difficulties accessing and processing information (Reid, 2016). Thus, there are modifications of definitions based on the interpretation of the current state of the understanding of dyslexia.

Reid (2016) indicates that despite the aim of a definition for dyslexia being to provide guidance, information, pointers for intervention and to develop an awareness of dyslexia, the wide variation of definitions can also cause confusion, provide little guidance and generalize a difficulty that can be very individual. The purpose of the definition can lead to a variety of criteria being included. A research definition may need to provide discrete and well-defined terms to outline the basis for selection of a sample of participants (Reid, 2016). Conversely, different definitions may be required to help parents, teachers and those with dyslexia understand the impact of dyslexia on learning (BDA, 2009). A definition may also be required to outline when resources may be allocated to support those with dyslexia. Definitions often include an outline of additional difficulties e.g. BDA, Dyslexia Scotland, DAI and IDA, which may imply that the outlook for those identified as having dyslexia is negative. It could be argued that a positive slant should be adopted so information, especially that aimed at parents and families, may include possible strengths identified in those with dyslexia in addition to weaknesses (Dyslexia Foundation of NZ, 2008).

2.7. Theories of causes of dyslexia

The fact that there is no universally accepted definition for dyslexia could be explained by the fact that there is no consensus of the causal factors of dyslexia. Many theories have been postulated with some gaining more acceptance and credibility than others. Possibilities that have been suggested as causes of the difficulties experienced by some with dyslexia include: phonological deficits (Castles & Friedmann, 2014; Melby-Lervåg et al., 2012); deficits in auditory processing (Giraud & Ramus, 2013); rapid automatized naming dysfunction (Lervåg & Hulme, 2009); difficulties with visual noise (Sperling et al., 2006); a cerebellar deficit (Nicolson et al., 2001; Stoodley & Stein, 2013); a deficit in temporal processing (Casini et al., 2017); a magnocellular deficit (Stein, 2014); and a visual attention deficit (Lobier & Valdois, 2015). Bosse et al. (2007) indicate that some research treats dyslexia as a unitary syndrome with a single underlying cause, whereas other studies find that dyslexia has diverse underlying deficits, which give rise to a spectrum of difficulties (Shovman & Ahissar, 2006).

Evans (2001, p7) asserts “all roads to reading start with the visual perception of the written word”. Stein (2014) states that even though reading problems are thought to be primarily due to phonological difficulties, the very essence of reading is visual. Every reader has to interpret and process visual symbols before they can translate them into sounds or access orthographic representations in their mental lexicon. Hence, even if a deficit in the visual perception of letters is not regarded as an underlying cause of dyslexia, it still has a role to play in enabling readers to access print at a basic level.

2.7a The Phonological Deficit Theory

The phonological deficit theory, the most widely accepted theory of dyslexia, postulates that a majority of readers with dyslexia have impairment in the representation, storage and retrieval of speech sounds (Ramus et al., 2003). Ramus (2014) asserts that it is the role of phonological deficits that set dyslexia apart from other reading difficulties and should be used to categorise dyslexia even though it can be argued that not all readers with dyslexia have a phonological deficit (Elliot & Grigorenko, 2014). As has already been discussed in section 2.3, learning to read includes making a connection between speech sounds and letter symbols, linking phonology and orthography. Young readers with dyslexia are generally poorer than controls at detecting, identifying and manipulating the onset of words e.g. finding the odd word out of tap – tot – mat – ten by identification of the initial /t/ ; or perceiving the rime of e.g. hat, cat, bat, as being the same but tag as being different (Cain, 2010). If a phonological deficit is present, then poor reading performance results (Vellutino et al., 2004). Prior to learning to read, if a child finds it difficult to distinguish sounds in verbally presented words then it could be predicted that they may have problems learning the alphabetic principle that letters represent sounds. Research by Melby-Levåg et al. (2012) and Snowling (2013) indicates that prior to reading instruction, difficulty with the conceptual understanding that spoken words consist of individual speech sounds and groups of sounds, i.e. poor phonological awareness, could be indicative of a future difficulty in learning to read when speech sounds need to be linked with printed letterforms.

2.7b The Magnocellular Deficit Theory

One of the theories of dyslexia that concentrates on visual processing and the visual system is the Magnocellular Deficit Theory. Stein and Walsh (1997) assert that the origin of dyslexic difficulties can be traced to the magnocellular system where reduced sensitivity creates difficulties in suppressing visual information. The magnocellular (large cell) visual pathway in the brain comprises of cells that detect contrast, motion and rapid changes in the visual field, whilst the parvocellular (small cell) pathway allows sensitivity to fine spatial detail. Research indicates that individuals with dyslexia have visual perception impairment stemming from disturbance in the magnocellular visual system (Cornelissen et al., 1995; Stein & Walsh, 1997). Retinal images persist longer than appropriate and this excess information means a reader with dyslexia may continue to perceive a word already read and link it to the next word. Stein (2014) outlines that the magnocellular system helps control eye movements and fixes the eyes onto a visual target. If there is a deficit in the magnocellular system then letters might appear to move around and cause visual confusion. These symptoms then interfere with learning to read, with readers perceiving letters and words incorrectly.

Evans (2001) asserts that 75% of those with dyslexia have a magnocellular deficit compared to 8% of the population as a whole. Support for this magnocellular deficit was found anatomically by Livingstone et al. (1991) by investigating brains post-mortem which revealed the magnocells in brains from people diagnosed with dyslexia were 30% smaller and more disorganised than in brains from those with no diagnosis of dyslexia.

One major criticism of the Magnocellular Deficit Theory is that it does not explain fully how phonological deficits can be apparent in readers with dyslexia. Hulme (1988) argued that if there were a direct relationship between reading difficulties and visual impairments, children with dyslexia would have more problems reading prose than single words, as a single word would not be subject to masking by an unsuppressed image from a previous fixation. This is rarely the case, with readers with dyslexia struggling with both single words and strings of words. Skottun (2005) presents an overview of studies where the magnocellular theory of dyslexia was investigated as central to explaining dyslexia. Of the 22 studies considered only 4 (18%) supported the theory; 11 provided positive evidence conflicting with the theory and 7 studies were inconclusive. In contrast, Stein (2014) presents more extensive evidence in support of the theory by outlining that over 90% of studies carried out in the previous ten years found evidence of impairment of the magnocellular system in at least some of their participants. Thus, he asserts that magnocellular deficits are a valid explanation of dyslexia.

Conversely, Ramus et al. (2003) state that a magnocellular visual deficit is often associated with a phonological deficit. This indicates that phonological difficulties and visual perception deficits are linked, although there is uncertainty as to how this occurs. If there is a connection, then neither the theory of phonological deficit nor visual deficit caused by weakness in the magnocellular system can provide a complete explanation for dyslexia

2.7c The Cerebellar Deficit Theory

Nicholson and Fawcett (2008) assert that a dysfunction of the cerebellum leads to a lack of fluency in skills that should be automatic. If a reader has not automatized the reading process then a conscious effort will be needed to apply grapheme-phoneme sound correspondences, which means that reading will require greater effort than for those for whom linking sounds and letters is automatic and effortless. It is suggested that in people with dyslexia the left cerebellum is larger but the number of neurons is less and more spread out, resulting in difficulties making connections (Nicolson et al., 2001). Stoodley and Stein (2013) argue that signs of cerebellar dysfunction are not found in all people with dyslexia. Their study of participants with cerebellar symptoms found that many did not have reading difficulties. They concluded that cerebellar deficit, whilst affecting some readers with dyslexia, was not an explanatory or causal effect.

2.7d Multiple Deficit Theory

There does not appear to be a single unifying theory that can explain the prevalence and behavioural aspects of dyslexia. A phonological deficit is agreed by many as being fundamental to the difficulties experienced by readers with dyslexia. However, other reasons for dyslexia are theorised to work in tandem with a phonological deficit and thus explain the broad spectrum of difficulties exhibited by those diagnosed with dyslexia. A Multiple Deficit Model (Pennington, 2006) allows for the heterogeneity of individuals with dyslexia to be explained by the presence of more than one deficit. A study by Carroll et al. (2015) produced results that support the view that reading difficulties have multiple causes. No single deficit appeared to be the main, underlying cause of reading difficulties. Predictors of dyslexia were poor print knowledge, i.e poor performance in letter-sound knowledge, sight word reading, and digit recognition; reduced phonological awareness, weak verbal short-term memory, and slow rapid naming. These findings support the Pennington (2006) model indicating that multiple deficits underlie dyslexia.

Building on the Multiple Deficit Model (Pennington, 2006), van Bergen et al. (2014) propose a model that incorporates an intergenerational aspect whereby it is theorised that the offspring of parents with dyslexia, or children with other close family members with reading difficulties may inherit genetic risk factors that could increase the incidence of dyslexia in these children. It is asserted that those identified with familial at risk factors may show cognitive deficits before reading instruction commences, indicated by impaired phonological awareness and reduced rapid naming skills.

Kelly and Phillips (2016) have devised an integrated causal model to explain causation with the interplay of a Phonological deficit, Magnocellular deficit and Cerebellar deficit at its core and assert that assessment of dyslexia should look for both single and multiple deficits. It is thought that between 30% and 50% of learners with dyslexia have multiple deficits (Georgiou et al., 2012; Pennington et al., 2012).

The interventions offered to children with dyslexia can be individualised to reflect the severity of dyslexia and the deficits identified. Interventions may be offered in small groups or as one-to-one support and usually take the form of an accelerated learning programme focusing on phonological awareness training and emphasizing phoneme-grapheme correspondences. Overlearning and repetition are also needed to support the acquisition of automaticity and reading fluency.

A phonological deficit indicates that the cause of dyslexia lies outside the visual system, which might imply that manipulation of font design would not address the fundamental issues faced by those with dyslexia when reading text. Thiessen et al. (2015) argue that typeface and cognitive load are connected. If a font designed to accommodate the needs of those with dyslexia were beneficial then it would allow individuals to spend more cognitive effort on retrieving phonological representations, which, according to Boets et al. (2013), are intact but difficult to access. A less legible typeface requires the brain to work harder to access visual pathways to identify graphemes. Thus, there is less residual capacity for linking orthography and phonology. If improving legibility decreases the effort needed to process text visually, cognitive load is reduced. This would provide increased cognitive capacity to access words from the mental lexicon or construct phoneme-grapheme correspondences, which would support the reading process. Van Someren (2013) indicated via a study of the specialist font Dyslexie that if letters are presented in a font in which the letters are easy to recognise, then the acquisition of sound-grapheme links is promoted. Furthermore, the study indicated that using Dyslexie font, made it easier for readers with dyslexia to recognise and interpret letters quickly and this led to improved automaticity, reading accuracy and reading speed. If these assertions are correct, then it could conceivably be the case that a change in font type could help address the underlying phonological deficits indirectly and impact the ability to decode and comprehend. If addressing legibility and then adapting it to the needs of the reader with traits of dyslexia enabled interventions that support a variety of deficits to be more effective, then this might reduce the level of reading difficulty experienced.

2.8 Summary

Reading starts with the eyes when visual processing is needed to perceive the written word. Linking graphemes to phonemes can then begin and written language can be decoded via the lexical or sub-lexical route, with comprehension then required for reading to occur. Reading progress is generally assessed by monitoring reading accuracy, reading rate and reading comprehension. Children whose reading does not develop as expected may also be assessed for dyslexia, but whilst there is no universally accepted definition or test for this, it is widely regarded as a reading difficulty that is underpinned by varying difficulties with phonological processing difficulties as a core deficit. Dyslexia is regarded as occurring on a spectrum from mild to severe and does not impact all of those affected in the same way. Individual intervention programmes to ameliorate difficulties are usually advised.

Chapter 3: Literature Review (Part 2):

Reading and Typography

As discussed in Chapter 2, reading depends initially on perceiving letters and then linking these written representations, graphemes, to spoken language in the form of phonemes, that comprise the words to be decoded. This chapter considers whether font design and the layout of print influence a reader's ability to identify letters and words and thus facilitate more accurate and fluent reading. It examines whether typography can make a passage of text more accessible to the developing reader, especially those who are struggling with the process of learning to read.

3.1 'Font' and 'Typeface' terminology[4]

For the sake of clarity, it is necessary to discuss the term 'font' and 'typeface'. To the layman 'font' indicates the design of a group of letters, which share a common style and compositional features. However, designers use the word 'typeface' to describe such a set of letters sharing a design and more precisely the word 'font' refers to characters from a specified 'typeface' of a particular size or weight. The word 'Typeface' originates from the times when all print was set out in frames with metal letters. A typeface comprised thousands of physical, metal blocks with each letter set out in relief to form the shaped surfaces that were inked by printers (Bigelow, 2019). Within the typeface i.e. the family of letters sharing a common design, each different character representing a letter, number or symbol, needed a specific, individual block

[4] For additional glossary of typography terms see appendix 1

for every size or weight. The ‘fonts’ were the characters of each particular size or weight and the ‘typeface’ referred to the design. A typeface will include many fonts within a common design e.g. Helvetica is a ‘typeface’ and Helvetica 10 pt bold is a ‘font’. A typeface “customarily has two hundred and twenty eight characters, including letters, accents, numerals, fractions, ligatures, commercial signs, and punctuation marks, ampersands, and peculiars, such as asterisks and daggers for footnotes” (Wilkinson, 2005, p. 57).

However, the advent of computers has bought with it its own lexicon and associated terminology. The word ‘font’ no longer just describes a block used to print a character with a specific design, size and weight but now has come to mean the entire collection of a named family of letters, numerals and symbols of whatever size or weight that is required (Opsteegh, 2010). When writing a document on most computer publishing programmes the user is asked to select a ‘font’ from a tab that includes design, size and weight. Thus ‘font’ has become a term recognised as covering all letters and symbols of a particular unifying design regardless of size or weight. Brownlee (2014) asserts that for most people, except designers, the words ‘typeface’ and ‘font’ are now usually used interchangeably, regardless of the terms’ historical roots. Conversely, Strizver (2014) argues that using the terms ‘typeface’ and ‘font’ as synonymous is erroneous, as it shows a misunderstanding of the terms. For this study ‘font’ and ‘typeface’ will be used synonymously, as the focus of the study is education and not graphic design, and the texts and passages used were reproduced using a computer, selecting ‘fonts’, ‘font sizes’ etc.

3.2 'Legibility' and 'Readability' terminology

It would also be helpful to address the terms 'legibility' and 'readability'. Both relate to how text is presented on a page but the difference between the meanings of the two is a matter of debate and opinion.

In a seminal work quoted by many font designers, Tracy (1986, p.31) claims that 'legibility' is the "quality of being decipherable and recognisable" and concerns the clarity of the single characters. 'Readability' is a measure of the length of time that text "can be read without strain or difficulty". This distinction is not always the same for every researcher. Strizver (2018) states that the word 'legibility' should be used when considering the design of typeface and the shape of letters or glyphs. A similar, somewhat narrow definition is also adopted by Bessemans (2016, p.21) who states "legibility is the ease with which visual symbols are decoded". In contrast 'readability' should be used when discussing how a font is arranged or typeset and should include elements such as size, case and spacing (Strizver, 2018). This distinction i.e. 'legibility' relating to the font design and 'readability' relating to layout is also used by Haley (2017); Dyson (2016); and Farley (2010). Unlike Tracy (1986) this distinction does not include factors other than typographical components.

In further contrast, Zhou et al. (2017) assert that 'legibility' should be defined by elements that are involved with perception e.g. clarity of individual characters, spacing, layout; and 'readability' is linked to elements of cognition e.g. comprehension and style of writing. This distinction is similar to that outlined by Luna (2018), who contends that 'legibility' concerns the identification of the letters when they are set together on the page, whereas 'readability' is a measure of how difficult a passage is to understand and includes sentence length, word frequency, multisyllabic words and syntax, contending that readability does not depend on typographic presentation. This definition also indicates that readability can vary depending on the educational level of the reader i.e. a young Key Stage 2 (KS2) reader may find a passage less readable than an older, KS2 reader because of the difficulty of the text, which is not dependent on visual elements.

In contrast to the above definitions, Beier (2009) asserts that for all except graphic designers, the two terms can be used interchangeably and for many they mean the same i.e. how clear the text is and how easy it is to understand. For the purposes of this study 'legibility' will be used to mean any aspect of typography that influences the perception of what has been produced on the page and 'readability' will include additional cognitive aspects, in line with Zhou et al., (2017) and Luna (2018).

3.3 Print: Legible but invisible

When text is published the font is usually chosen carefully by designers to give the best reading experience. The print should convey the information from the text whilst not directly intruding into the reader's thoughts. Cullen (2012, p7) outlines that the craft of the typography designer is "to make language visible" but for "the type to go unnoticed." This concept of the invisibility of print was outlined by Warde (1956) using the metaphor of a crystal goblet, which represented the print, with the red wine contained within the goblet being the meaning of what is written. The crystal goblet is beautiful, perfect and crafted by an expert but allows the contents to be viewed and enjoyed unhindered through the transparent glass. Thus, print should enhance and present thoughts without imposing or distracting attention. This premise has since been adopted by many typographers, who assert that the reader should not notice that the type is there (Beier, 2012; Luna, 2018; Lupton, 2014; Shafig, 2015). Breuninger (2019) emphasises this by stating that few appreciate good print design, but many notice poor print design. This signifies that just because type often goes unnoticed it does not mean that it is unimportant or has no influence on the reader.

Strom (2003, p.9) argues that reading difficulties can be incorporated into Warde's (1956) metaphor by describing dyslexia as a "fog that complicates the transparency of the goblet." This may imply that the reader with dyslexia has to pay more attention to the print and decode consciously rather than is the case with typical readers, who access the meaning automatically without needing to concentrate on the print, which is to all intent 'invisible'. The 'goblet' for the reader with dyslexia is not transparent but clouded so that the printed word presents a barrier to accessing meaning immediately, and hence, the fog needs to be cleared to see the contents. It is possible that by manipulating legibility, i.e. presenting text in a different font, the 'glass' may become less foggy and less noticeable. In other words, the reader with dyslexia does not have to work so hard to decode the words so that the overall task of reading becomes less arduous. If a particular font allows reading to become more automatized and the meaning accessed more easily through a less 'foggy' glass, then the impact of dyslexia becomes less intrusive.

3.4 Letter identification

To assess whether a font has the effect of increasing legibility, it is necessary to consider the typographical features that influence legibility. However, even before ascertaining whether a text is legible, it is necessary to consider how a reader identifies letters, which are the basis of legibility. The neurological circuits involved in print perception have already been discussed (section 2.2b) but the way in which letters are identified needs to be deliberated. There are two main theories on the process of letterform identification, namely that whole letters are read as a visual template or letters are read as a combination of features (Herrmann, 2011).

The first theory is that the brain has a stored template of all letterforms and as a new letter is perceived, the brain scrolls through the existing templates to find the best match (Hunziker, 1998). The main difficulty with this theory is that readers are capable of recognising a huge range of typefaces, together with individual handwriting styles, and it is difficult to believe that all the possibilities are stored as individual templates.

The second theory is that instead of perceiving whole characters the brain decodes specific features contained within letter shapes. These disparate features are combined until an identification of a letter is made (Beier, 2012). Changizi et al. (2006) analysed 96 writing systems to identify which shapes and design elements are commonly contained within letterforms. They then determined which shapes that are found within written alphabets were commonly observed in the environment. For example, two edges forming a junction corresponds to the Latin/Roman alphabet letters ‘L’ or ‘T’, with more complex junctions equating to ‘F’ ‘Y’ ‘E’ ‘X’. Curves and holes also form elements of letters such as ‘O’ ‘Q’ ‘P’ ‘D’. It was found that most common letter shapes match the shapes that people encounter daily in the visual world (Changizi et al, 2006). These natural shapes, referred to as visual invariants by Daeharne (2009), have been integrated into our visual recognition system and so were capitalized on when alphabets were devised. As has already been discussed in section 2.1, reading is not an innate skill, so utilizing shapes that the brain identifies naturally and easily uses existing mental skills to help to achieve a literate human.

Analysing how readers can detect the location of specific characters supports the theory that letters are identified by certain elements rather than by the entire shape. Neisser (1967) details that participants found it easier to locate the letter ‘Z’ among unrelated characters ODUZGQRC than a group of related characters, IVMZXEW. To support the template theory readers would not be expected to find the location of the target letter any more easily in either condition of unrelated or related letters. The whole target letter would be identified regardless of surrounding letters. If a certain element of the target letter were sought, then it would be more distinguishable among unrelated characters that do not include that element than related characters that also have that feature.

When learning to read children learn to distinguish critical features. Early readers have been shown to focus unpredictably and randomly but children with up to three years reading experience tend to look at the parts of the letter that carries the most information, i.e. the parts of a letterform that make a letter distinguishable (Willingham, 2017). These identifying elements that inform the reader what the letter is can include horizontal lines, vertical lines, diagonal lines, corners, semi-circles etc. Identifying letters by features not as a whole allows letters to be written in different fonts and still be recognisable.

As shown in Table 3.1, Garamond, an old, traditional style font and OpenDyslexic, a modern, specialist font have the letters b/d, q/p, n/u, formed so that they are not flipped or rotated versions of each other, which is a design aspect found in certain other fonts e.g. Arial.

Table 3.1 Comparison of fonts with unique letterforms v flipped or rotated forms

Garamond	b d, q p, n u
OpenDyslexic	b d, q p, n u
Arial	b d, q p, n u

Garamond is designed using varied shaped serifs to distinguish letterforms and OpenDyslexic has varied letter shape, weight, and counter size to achieve distinctiveness. The concept of each letter having unique features was considered by Herrmann (2011), who suggested that distinct letters may make a particular font more comfortable to read but a unique design for every character is not necessary as letters are not read as a whole but, as discussed above, as a series of features. If a reader is detecting parts of a letter, then other elements that may be similar in a reversed letter may not be influencing identification. However as has already been considered, developing readers focus randomly on letterforms so may not automatically focus on the distinctive elements in the way that an adult reader might. It is conceivable that a child with traits of dyslexia would not instinctively focus on distinguishable parts of a letter but look at all elements of its shape indiscriminately. A unique letterform may support them in ways that a competent reader might not need. Thus, legibility for some may rely on uniqueness whereas for others orientation may be sufficient.

Fiset et al., (2008) analysing letters presented in Arial, identified letter features that facilitated recognition. These were mainly stroke terminations e.g. the two terminal ends of a letter ‘C’ or the bottom curve of a letter ‘j’, or horizontal strokes e.g. the cross of a ‘t’. They also found that participants focussed attention on regions of a letter that could be confused with another, e.g. to distinguish ‘c’ from ‘o’ the point of focus needs to be the right-hand section that is open on the ‘c’ but is closed on the ‘o’. The other parts ‘c’ and ‘o’ are identical so don’t need close attention to enable the reader to draw a distinction (Fiset et al., 2008).

Despite each character being a combination of elements that make it individual, some letters can be easily confused with others. Beier (2012) identifies groups of problematic characters that are easily confused. These letters can be grouped to show commonalities that reduce individuality and thus lead to possible confusion (Table 3.2). Beier (2012) also indicates that the misreading of letters is also impacted by font design. In a meta-analysis of studies, it was found that, for example, the letters ‘l’ and ‘i’ are the most misread pairs when printed in the font Courier but by comparison rank fifteenth on the list of most confused letter pairs when the characters are printed in Futura (Beier, 2012).

Table 3.2 Pattern of recurrent misreading of letters (Beier 2012)

	Confusable letters	Commonalities
Lower case	e c a s n u o	x height size, no ascenders or descenders Mixture of curved and straight line
	i j l t f	Narrow, small width, single vertical stroke
Upper case	O Q D C G	Round shapes
	V Y M V K X	Diagonal shapes
	T I J L	Vertical stem shapes
	F B P E T H	Vertical and horizontal strokes
	H N M	2 vertical stems

The role of fonts in aiding or confounding clarity is addressed by Bohm (2019), who asserts that a lowercase l (L) can be confused with a number 1 or a capital **I** (i) especially in fonts that do not have serifs. This is illustrated by comparing the fonts Bucko and Sassoon Infant (Table 3.3)

Table 3.3. Comparison of Bucko and Sassoon Infant fonts to show how characters may be confused (Bohm, 2019)

Bucko:	Henry III	Henry Ill	Henry III
Sassoon Infant	Henry III	Henry Ill	Henry 111

It can be seen that the distinction between what is written in the three versions in Sassoon Infant is quite clear. However, Bucko makes distinguishing between an uppercase I (I) and lowercase l (l) and a number 1 difficult. The upper case I (I) is fractionally shorter than the lowercase l (l) but these require close inspection. The lowercase l (l) and the number 1 are indistinguishable. Thus, for this example, Bucko could be regarded as having poor legibility because of the lack of individuality in some letterforms.

3.5 Aspects of font legibility

Many combined elements of typography work together to form a page of legible text. It has been seen that some fonts seem to be more legible than others so it is now necessary to consider the elements of font design that make a font legible. Zorzi et al. (2012a, p.1.) assert that a valid way to complement the effectiveness of reading interventions is to focus additionally on “making reading material more accessible by manipulating the physical properties of print e.g. print size, font type etc.” Thus, the elements that combine to improve legibility should be modified to allow the reader the most effective choice of font design to support an optimum reading performance. The way that this might be achieved is not straightforward. The effects that individual elements of font design have on legibility are difficult to unravel. Even if aspects of legibility can be isolated there is still the human element of each reader reacting differently to different stimuli and responding to texts individually. Researchers have investigated the ways that many elements of font design can potentially impact reading performance. Not all of the individual elements influencing font legibility will be considered here. The focus will be on the constituents of font design that are most relevant to this study.

3.5a Font size and legibility

An obvious and important element of legibility is font size, which is measured in points (pts). There are historical and international differences, plus metal block print and digital variations of a point size, but it is usually recognised that there are 72 pts to an inch, with a point thus measuring 0.0138 inches or 0.351mm (Garfield, 2011). The actual size of the font is not defined when setting the point value but rather the body height, which in manual typesetting is the height of the lead type or metal block on which the actual font face is moulded (see Figure 3.1). Hence, it is historically the metal type block and not the printed letter that has a vertical height of a specified number of points. Font designers are free to decide, even when designing in the digital era, how to use the body height to create a balance between each character and the non-printing, blank space around it. Depending on how the available space is used, the measurable font dimensions may vary, so that two different fonts set in the same point size can appear similar in size or look very different (De Soto, 2014).

Figure 3.1 How body height and point size are related (Breuninger, 2019)

In children's texts, the size of the characters is usually large initially and decreases in size as the age of the child at which the book is targeted increases. Hughes and Wilkins (2000) studied 120 children who took a reading test with the print design based on the print size of a series of reading scheme books that had been published with font sizes becoming smaller as the difficulty of the text increased. They found a significant decrease in the reading speed of the children aged 5 to 7 as the font sizes decreased, whereas participants aged 8 to 11 showed no significant effect. Children in both age groups made more reading errors when using the smaller font sizes. They conclude that presenting children with a larger font size could increase reading speed and accuracy. The tests that were used in their study did not include meaningful texts that measured comprehension but in a further study (Wilkins et al., 2009) investigated speed of comprehension when font size was increased. The font used was Arial and texts were comprehended more quickly when presented in a larger size. There was also a marginal increase in accuracy. O'Brien et al. (2005) found that for young readers with dyslexia to achieve their maximum reading rate, a font that is approximately 32% larger than that needed by typically developing readers is required. This indicated that although a large font size did not completely enable readers with dyslexia to read as quickly as their typically developing peers, it did allow them to read at the maximum speed that they were able. Thus, these studies imply that reading tests that are produced with the font size reducing as the difficulty level increases may penalise some readers who could score more highly if the font size was increased or at least, not decreased.

Katzir et al. (2013) also investigated the effect of font size on the comprehension of young readers. Participants were in two groups comprising 45 from grade 2 with a mean age of 7.5 and 45 from grade 5 with a mean age of 10.5. The text used was presented in a variety of font sizes and line lengths with spacing manipulated. For the younger group. the results showed that the smaller font size together with increased line length lead to significantly lower comprehension scores, which triangulates with the findings of Wilkins et al. (2009). However, the older group achieved higher comprehension scores when using a smaller font size. Katzir et al. (2013) suggest that this is because the smaller font lead to deeper processing for the more competent readers which impacted on the retention of what they read, thus leading to higher comprehension scores.

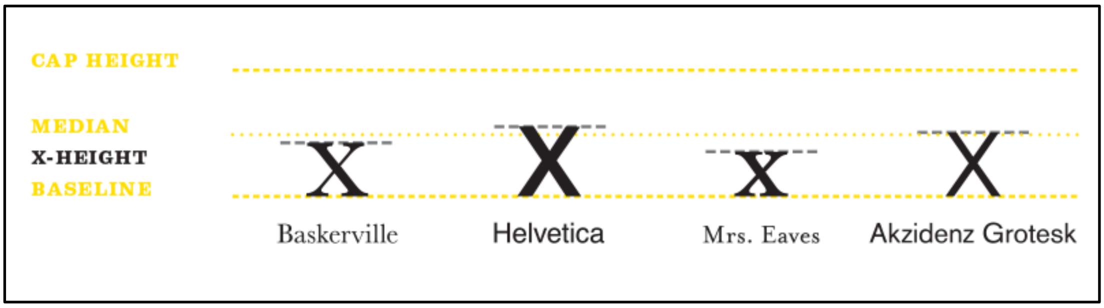
The variation in findings of these studies shows that the impact of font size on developing readers is not consistent and is possibly working in conjunction with other variables such as font design, line length or spacing plus characteristics of the participants such as age, stage of reading development, reading automaticity and reading accuracy.

Font size and the perception of fonts on the ease of reading was studied by Bernard et al. (2001). Twenty-seven developing readers aged 9 to 12, who read coherent text presented in 4 fonts in both 14pt and 12pt were asked to answer a perception of readability questionnaire with answers ranked on a Likert scale. The participants perceived that they read more accurately and quickly with the larger font size. The study considered font size in terms of the readers' perception of legibility, which is subjective. Zender (2019) suggests that a legibility scale should be devised that includes perceptual as well as measurable features to score the letterforms within a font and then make a composite score for each font family to compare legibility consistently and accurately. This would be less susceptible to idiosyncratic perceptions of font legibility. However, this has yet to be devised and standardised.

3.5b X-height and font legibility

Closely aligned to font size is x-height. The x-height of a letter is the distance between the base line and the mean line, which is the top of a lowercase letter (figure 3.1). An ascender rises above the mean line and a descender below the base line (Garfield, 2011). Beier and Dyson (2014) explain that some fonts of the same point size have varying x-heights and for the x-heights in these fonts to appear equal the point size has to be increased. This difference can be seen in Figure 3.2, where the variation in x- heights of fonts of equal point size is clearly illustrated (Stinson & Elnar, 2016). A font with a point size of 10pt but a large x-height may look the same as a font with a small x-height presented in point size 12pt (Zhang, 2006). When a small point size is used, a font with a large x-height is more legible than a font with a small x-height (Loyd, 2013). However, if the size of the x-height compromises the amount of space left for the ascenders and descenders then words can become illegible if, for example, 'n' and 'h', become indistinguishable.

Figure 3.2. Variation of x-height in equal point sized fonts (Stinson and Elnar, 2016)



Beier and Dyson (2014) assert that in any form of research that concerns typography, equating x-heights rather than point size is vital. If point size is controlled for rather than x-height then the appearance of letters can be as if they are of different sizes. If x-heights are matched then a comparison is more valid than if point sizes are equal. Herrmann (2011) states that any study on legibility that compares typeface of the same point size but does not emphasise x-height equivalence rather than point size will have flawed results. Thus, the effect of x-height is a factor that should not be overlooked.

3.5c Serifs and legibility

A serif is the small finishing stroke added as embellishment to the form of a letter character (Garfield, 2011). Fonts that include these elements are known as serif fonts and those that do not are sans serif fonts. In typography there is a belief that serifs increase letter discrimination, with characters being heterogeneous due to the presence of serifs (Bessmann, 2016). Thus it is surmised that a text is more legible if presented in a serif font, as letters will not be confused. The end strokes of a serif font are also argued to enhance the ability of readers to trace a line of type clearly and that may lead to faster and more efficient reading (Arditi & Cho, 2005).

Conversely, serifs may act as a form of visual noise and detract from overall legibility (Moret-Tatay & Perea, 2011). This visual noise hypothesis may be valid if the letterform is viewed as a template as discussed above (section 3.4), but if only certain elements within the letterform are identified then it is conceivable that the reader may ignore the presence of serifs. The majority of research comparing serif and sans serif fonts has found little difference in reading performance (Beier & Dyson, 2014). Beier and Oderkerk (2019) assert that the comparison of the effect of serifs by just using two different fonts would not be valid. For an accurate appraisal of the influence of serifs, the fonts selected must be controlled so that they vary in design by only the presence or absence of serifs, and do not have in addition other variables such as different stroke widths, weights or x-heights. For example, sans serif Helvetica and serif font Times New Roman are not comparable, as the two typeface styles vary in other aspects not just the presence or absence of serifs (Beier & Oderkerk, 2019). Some published studies have compared serif and sans serif fonts but have not controlled for other confounding factors(Beymer et al., 2008; Josephson, 2008). However, Morris et al. (2002) used two fonts from the same family, Lucida and Lucida Sans, so that all elements were controlled for. They found that serifs reduce legibility at small sizes but this was not significant when viewed at or above point size 16. Moret-Tatay and Perea (2011) also used Lucida Bright and Lucida Sans, in a word discrimination task and found no benefit for words written in a serif font. They found a slight increase in identification times for the sans serif version. They suggest that this supports the use of sans serif fonts on traffic signage where reaction speed is important, but when reading blocks of text for understanding there is no significant benefit between either serif or sans serif fonts. Arditi and Cho (2005) constructed nine lower case fonts that varied only by serif size and spacing conditions resulting from the size of the serifs. The smallest size font had the effect of increasing legibility due to increased spacing, but the presence or absence of serifs had no effect on reading speeds.

All of the studies concerning serifs considered above had adult participants. The effect of serif or sans serif fonts on beginner and developing readers should be considered as serifs may influence children’s reading performance differently. The common view, held by many in education, is that children’s texts should have sans serif fonts. Simple sans serif letterforms are assumed to be better for beginner readers who might be confused by unnecessary serifs (Sassoon.1993). Walker and Reynolds (2003) used miscue analysis to analyse reading errors when children read in Gill Sans and Gill Schoolbook, which are sans serif fonts, plus Century and Century Educational, which are fonts with serifs. They found no difference in reading performance in either serif or sans serif fonts. Thus it can be surmised that as children see many fonts and graphics from an early age, even as pre-readers, they learn to recognise letters with a wide variety of shapes but can distinguish the crucial, basic characteristics of each letter and are thus able to identify letter forms with additional attributes such as serifs.

The choice of serif or sans serif font to support those with dyslexia has also been debated. In the 1990's the British Dyslexia Association (BDA) produced books that targeted readers who had dyslexia in Times New Roman, which is a font with serifs (Smythe, 2010). At this time it was thought that serif fonts led the eye from one letter to another in the manner of cursive fonts and were thus easier to read. From 2000, BDA publications were produced using the sans serif 'Arial' font, which was used until 2007 when the font used for the Dyslexia handbook was changed to Futura, which is also a sans serif font but with a rounded shape '**ɑ**'. This could have been because it was surmised that this made reading easier for those with dyslexia. Smythe (2010) indicates that although sans serif fonts are now always recommended for use with readers with dyslexia this recommendation is not as a result of widespread consultation or detailed research but due to a "feeling" that BDA members would prefer a sans serif font (Smythe, 2010, p.121). It seems that at present the trend is for books aimed at young readers or those with dyslexia to be published in sans serif fonts. Fonts with serifs are still used habitually in textbooks, information books and newspapers.

3.6 Layout and legibility

Thus far legibility has been considered in terms of the attributes that comprise letterforms. Legibility of a whole text is not just due to letters but also the combination of letters to make words and the layout of words on the page. Some factors that are concerned with the clarity of words and page layout are now considered.

3.6a Line length and legibility

In text presented at standard book size, it is recommended that line length should average 50-70 characters per line (cpl) including spaces and punctuation (Luna, 2018). Newspaper columns should average 35-45 cpl (Mileta, 2014). These line lengths will vary according to page size, font size and in the case of newspapers or magazines the number of columns presented on the page (Briem, 2002).
The layout of printed text and consequently the line length is fixed at the time of production but on screen the layout of a text can vary. Line length may change as a function of the default settings of the hardware platforms; different software versions; or the device used e.g. layout of the same text can vary when viewed on a smartphone, tablet, laptop etc. (Thomson et al., 2018). Altering the font size also automatically alters the number of characters on a line (Luna, 2018). This means that designing an optimum line length for a text to be presented on screen is difficult, as aspects outside the designer's or even reader's control can determine line length.

Davidow (2002, p.22) asserts that when designing a text layout the “visually appealing should not be confused with the readable”. Thus, it is necessary to consider how line length impacts legibility in terms of reading performance rather than how aesthetically pleasing the page looks.

Over the last few decades, research into the effect of line length on reading has focussed on reading on screen. This reflects the growing use of screens and electronic reading devices to read texts originally or concurrently presented as print on paper. Text is also easy to manipulate on screen so variables can be controlled for research purposes. Research does not however confirm a particular line length as being most legible. Dyson and Haselgrove (2001) found that a medium line length of 55 cpl was most effective to support comprehension and reading speed, whereas Shaikh (2005) found lines of 95cpl supported faster reading speed but had no effect on comprehension. In contrast, the participants in Rello and Baeza-Yates’ (2017) eye tracking study preferred a shorter line length of 45 cpl but line length was shown to have no effect on eye fixations.

Two studies considered the impact of line length on the reading performance of high school-aged readers with dyslexia. In an eye tracking study, Schneps et al. (2013a) used an Apple iPad to present texts in lines averaging 11.6 words per line (67.2cpl) and an Apple iPod with lines averaging 2.19 words (12.7 cpl). All passages were in 32pt Georgia font in either normal or expanded spacing and were 208 words long. The iPad displayed the full text on one screen in landscape mode and the iPod required 12 pages to display each text. Data from 25 participants were analysed. Findings indicated that the use of the iPod and thus reading using shorter line lengths improved reading speed by 27%, reduced the number of fixations by 11% and reduced the number of regressive saccades by half (Schneps et al., 2013a). In the second study (Schneps et al., 2013b) compared text presented as black print on white paper in lines of average 13.94 words per line printed in Times New Roman, 14pt, and text presented electronically on an iPod with 3.4 words per line set in Times New Roman 42pt and scrolled as a continuous text. Findings also showed a benefit from reading very short lines on the iPod but only for certain subsets of participants. Those who had difficulties with decoding phonemes and sight words were found to read with an increased rate using the iPod. The group of readers who had been identified as having poor performance on Visual Attention Span tests showed improved comprehension when reading on the iPod. This indicates that the short lines may be influencing the processes that support visual attention and supporting deficits that this particular subset of readers with dyslexia has to overcome. Conversely, the group of participants that scored highly on the Visual Attention Span tests, scored more highly on the comprehension tests when reading using print on paper and thus possibly benefitted from reading text with longer lines. It is clear that as dyslexia impacts readers in individual ways any modification to help address difficulties also impacts readers differently. It is only possible to say that short lines help some readers with dyslexia under certain conditions. Line length supports legibility and subsequent reading performance for some struggling readers but not all.

Walker (2001) indicates that line length should depend on the nature of the information presented and who the text is intended for i.e. the parameters should be different for a book designed for developing readers and one intended for university undergraduates. Books aimed at young readers frequently have lines that equate to sentences with each new sentence beginning on a new line. Even if sentences cross line boundaries they rarely cross beyond a double page. In this way a reader does not have to turn a page in the middle of a sentence. It is also argued (Raban,1982) that for children breaking a line between phrases causes less disruption to fluent reading than breaking text according to line length. Breaking lines according to sense may help the beginner reader keep track of the meaning of the narrative. Conversely, some children's books ignore typographical conventions completely and have lines that meander and cross the page to link illustrations, indicating that children have a broad tolerance of text layout variations (Walker, 2005).

3.6b Spacing and legibility

The spacing of the letters, words and lines on the page can influence legibility hugely. Walker et al. (2018) state that elements that affect text can be categorized in two ways; intrinsic variables that are part of the font, such as shape and other designed characteristics, and extrinsic factors, which are features that can be changed e.g. the space around the letters or words. Thus it is not just characteristics of font design that can influence legibility but the surrounding extrinsic features that may be exerting an effect on reading performance. Two main aspects of spacing can be identified, inter-letter spacing and line spacing, both considered below.

i) Letter and word spacing

In typography, inter-letter spacing is often called tracking and refers to the amount of white space between two letters. Unlike kerning, which is the specific adjustment of spacing between two particular characters, inter-letter spacing generally applies to the overall spacing between all neighbouring characters and affects the character density of a line or a block of text. The size of inter-letter spacing affects text legibility as it manipulates the amount of lateral interference among neighbouring characters. Inter-letter spacing also affects inter-word spacing as the font selection or browser generally sets the gap between words as a default. There is variability in the amount of default inter-word spacing used across different fonts (Slattery et al., 2016). For this section, inter-word and inter-letter spacing will be considered in tandem.

Letter spacing, and its effect on reading performance, has been studied widely. A Boolean search of the library catalogue of ‘letter spacing’ and ‘reading performance’ provided over forty thousand references. Thus, this section will focus on the effect of letter spacing on the reading performance of children, particularly those with dyslexia. Table 3.4 shows a comparison of research studies specifically chosen as the participants were all children. Five of the studies compare results from readers with dyslexia and typically developing control groups. The focus of all of the studies is letter and word spacing rather than specific font designs. Studies that consider spacing alongside fonts specifically designed for those with dyslexia will be discussed in section 3.9 below.

Taken overall, the findings on letter spacing are inconclusive. This could be due to the variation in methodologies, languages, or of course the fact that every study had different participants who were all individuals with unique characteristics.

Table 3.4 Review of research on the effect of spacing (child participants)

Authors	Language	Participants	Task	Font	Conditions	Findings
Hughes and Wilkins (2002)	English	200 children age 6-11 years	Visual Acuity test: single letters. Word and lines Reading Rate test: randomly ordered common words, no semantic meaning Printed on paper	Geneva x-height 5mm or 5.8mm	VA: 4 x-height, 5 letter spacing; 5 line spacing RR: 2 x-height; default spacing plus word spacing increased by 5% and line by 10%	Increased spacing improved reading performance.
Reynolds and Walker (2004)	English	24 children age 5-7 years	Read continuous, meaningful text. Printed on paper Measured reading rate and miscue analysis of errors	Century Educational 19 pt	Letter spacing: tight (50% normal), normal, wide (170%), very wide (250%) Word spacing: tight(-20), normal, wide (+15), very wide (+25)	No significant benefit from spacing that was tighter or wider than default
Perea et al., (2012)	Spanish	18 readers with dyslexia(DR) age 11-13 years 20 normally developing readers (NR) age 9-10 years	Read continuous, meaningful text. Read on screen Measured reading time and accuracy. Comprehension test	Times New Roman 14 pt	Spacing: Default and +1.2	DR: wider spacing: significant benefit to accuracy, reading times and comprehension scores NR: no significant difference in accuracy, reading time or comprehension scores.
Zorzi et al., (2012a)	Italian/French	74 children with dyslexia (34 Italian / 40 French), age 8-14 years 30 control group, Italian, mean age 7.8 years	Read short, meaningful, but unrelated sentences. Printed on paper. Measured accuracy (no. of errors) and reading speed (syllables per second)	Times New Roman 14pt	Spacing: Default and +2.5	Dyslexia group: wider spacing: significantly few errors; significantly faster speed. Control group: wider spacing: fewer errors, faster speed but <u>not</u> significant
Van den Boer and Hakvoort (2015)	Dutch	197 children: 105 mean age 7y 10 m (grade 2). 92 mean age 9y 11 m (grade 4). Includes 29 poor readers (15 G2/14 G4)	Word naming fluency. Read 6 lists of 24 mono- or bi-syllabic words. Read on screen, presented singly Measured fluency; accuracy, speed	Times New Roman 14pt	Spacing: -0.5, default, +0.5, +1, +1.5	No benefit of increased letter spacing for either typical or poor readers. Reduced fluency when spacing smaller than default
Hakvoort et al. (2017)	Dutch	Exp.1: 60 children, 30 dyslexia, 30 control, (mean age 9y 11m) Exp.2: 189 children, mean age 9y 3 m	Read sentences. Print on paper. Measured accuracy and reading rate Read sentences on screen, either presented as whole, or word by word. Measured accuracy and reading rate	Times new Roman 14 pt	Spacing; Default and +2.5 Spacing: Default and +1.2	Reading speed unaffected. Accuracy improved in spaced condition for both groups. No effect when presented word-by-word. Increased accuracy in spaced, whole sentence presentation

With regard to reading speed, in these studies spacing appears to have no significant effect of typically developing readers and only the studies by Zorzi et al. (2012a) and Perea et al. (2012) demonstrate a beneficial effect of wider spacing on reading rate for readers with dyslexia. The study by Hakvoort et al. (2017) which aimed to replicate the study by Zorzi et al. (2012a), did not in fact find that reading speed was improved by wider spacing but their findings showed reading rates unaffected.

Accuracy of reading, considered in most instances as the number of decoding errors, also showed inconclusive findings. Zorzi et al. (2012a) indicated that fewer errors were made when reading using a wider spaced text, with results being significant for the participants with dyslexia but not statistically significant for the control group even though a reduction in errors was recorded. Skottun and Skoyles (2012) critiqued the study by Zorzi et al. (2012a) and suggested that the text presented to the control group was not at a sufficiently challenging level to allow for errors to be made at significantly different levels. Most of the participants made very few reading errors when reading using either spacing condition. Thus, they argue it was not possible to demonstrate that the findings for the control group were valid. Zorzi et al. (2012b) replied that in their opinion typically developing readers may benefit from wider spacing in the same way that readers with dyslexia were found to do.

As already indicated, Hakvoort et al. (2017) aimed to replicate the findings of Zorzi et al. (2012a). They endeavoured to ensure that the ceiling effect on reading errors for typically developing readers was avoided by using text of an appropriately difficult level that allowed room for improvement in reading accuracy (Hakvoort et al., 2017).

However, even amongst the six papers reviewed here, there is not a consensus regarding reading accuracy and letter spacing. Wider spacing was shown to be of no benefit to reading accuracy by Reynolds and Walker (2004), and Van den Boer and Hakvoort (2015), but to benefit the reading accuracy of all readers by Hakvoort et al. (2017). Readers with dyslexia were found to benefit from wider spacing by Perea et al. (2012) and Zorzi et al. (2012a).

It is apparent from the studies described that there is not one particular letter spacing that sets a 'gold standard', able to support the reading performance of all children in all reading conditions. The research overview indicates that default or slightly wider than default spacing is more likely to provide the best spacing conditions for young readers, with a reduction in the default spacing being unlikely to help reading performance.

Letter spacing is also linked with the phenomenon of crowding, which is when identification of a target letter is impaired due to the presence of neighbouring elements in the peripheral field of vision (Martelli et al., 2009). Words surrounded by other words, or letters surrounded by other letters are more difficult to identify than isolated words or letters. The effect is thought to be more pronounced for readers with dyslexia than others (Spinelli et al., 2002). It is possible that by expanding the spacing of a text the effect of peripheral letters or symbols that distort perception and cause the phenomenon of crowding may be reduced. The six studies considered above did not specifically address crowding but a variety of research tends to suggest that adjusting text by manipulating spacing to reduce crowding could be beneficial for both children and adults with dyslexia (Joo et al., 2018; Martelli et al., 2009; Moll & Jones, 2013; Pelli et al., 2016; Schneps et al., 2013b; Spinelli et al., 2002).

ii) Line Spacing

Line spacing, or leading, so named because of the strips of lead placed between each line to separate them by typesetters, is the vertical space between lines. Line spacing is commonly measured as a percentage of font size, conventionally for published text set between 130% to 150% (Clarke, 2018). Increased white space between lines, set at up to a maximum of 200%, can help separate descenders and ascenders that might otherwise merge into each other (Smythe, 2010). Setting lines too close, at 100% or less, can make them inaccessible for many, especially poor readers, children or the visually impaired, as lines of letters lose their separation and distinguishability (Clarke, 2018). If leading is set at over 250% the spacing is such that many readers find it difficult to move smoothly from one line to the next, which hinders reading speed and concentration (Luna, 2018).

Strizver (2011) suggests that children's books should have wide leading so that children can follow easily from one line to the next. However, Walker et al. (2018) outline that no optimum inter-linear spacing has been identified specifically for children's books. This is supported by a study by Reynolds et al. (2006) that compared four different line spacing layouts, with the help of 24 five to seven year olds. Reading performance was assessed by miscue analysis and the participants were then questioned about the varying page layouts. It was found that reading performance was not affected significantly by the four line spacing versions but the participants noticed the variations of layout and a majority (16/24) preferred normal or wide line spacing. The spacing offered as normal reflected the spacing of most children's books, which is generally wider than standard text not aimed specifically at children. Hence, the preference erred on the side of generous inter-linear spacing rather than narrow. These results indicate that whilst performance may not be affected, layout could influence motivation to read by line spacing adding to a book's appeal.

3.7 Font and Comprehension

Text readability as defined by Zhou (2017) and Luna (2018) (section 3.2) incorporates a cognitive component, so the impact of font design and text layout on comprehension should be considered alongside legibility. Franzen et al. (2019) investigated the relationship between font and reading comprehension and found that presenting text in OpenDyslexic rather than Times New Roman lead to an improvement in reading comprehension scores for both their participants with dyslexia and to a lesser but still significant extent for the control group. The participants in the study were adults, so would be experienced readers, with the participants with dyslexia conceivably having developed compensatory strategies over time. Many studies that address font and comprehension have adult, undergraduate participants, who would be very proficient readers with a level of comprehension skills sufficient to enter university (Diemand-Yauman et al., 2011; Dressler, 2019; Rummer et al., 2016). These are arguably not comparable with studies with participants that are children and as such still developing reading skills.

A study by French et al. (2013), with participants aged 13-16 years, focused on whether presenting educational materials in a hard-to-read font led to better recall of information. They argued that by making information harder to read by changing font, improved retention and learning was achieved. Participants with a cross-section of abilities were presented with a power-point slide comprising 60 words about a fictitious star to be read within 90 seconds in silence. A group of 154 participants were shown the facts in hard-to-read Monotype Corvisa font and 121 control group participants read the facts in Arial. After 35 minutes they were given 7 multiple choice questions to test recall. Results showed that all groups of students, including those identified as having dyslexia, demonstrated a significant increase in recall and retention when they read using the disfluent, hard-to-read font. The fact that the students read only 60 words leads to questioning whether younger readers, especially those with dyslexia, would be overwhelmed if a longer text were presented in a hard-to-read text and tire more easily. The increased effort may outweigh any benefit of the increased processing leading to improved recall. This study also tested recall after a time lapse rather than comprehension immediately after reading so placed more emphasis on memory retention than comprehension.

The results coincide with those of Diemand-Yauman et al. (2011) who found that a disfluent font lead to improved recall but are contrary to Rummer et al. (2016) who repeated the Diemand-Yauman et al. (2011) study but found that there was no effect of font. Conversely, Dressler (2019) found that using Times New Roman font rather than a hard-to-read font led to improved comprehension scores, especially when the participants were allowed unlimited time to read and respond. These studies were all methodologically different and do not show a consistent impact of font on comprehension when tested by assessing recall. None of these studies reported aspects of comprehension such as inference but relied on recall of facts.

3.8 Font design for children and young readers

It has been indicated by Sassoon (1993) that some young readers are able to read some passages of printed text but unable to read others, even though the vocabulary is at the same level. This could be due to the font being different and possibly eroding the identity and distinction of some letters and thus hindering decoding for some children. This indicates that the choice of font for children’s books may affect their ability to read a book. This premise is supported by Bessemans (2016), who makes the point that as details of font design may influence performance during reading, all fonts should not be regarded as equally suitable for children’s books. As visual details and design affect legibility, and the effect of font design on reading performance varies, consideration of the optimum fonts for children is valid.

Most reading material for children, especially that used in schools, is in printed form (Woods et al., 2005). Sassoon (1993) asserts that many children’s books are published with the layout and font being chosen for sales appeal and aesthetics rather than legibility for the child. The criteria for making a page attractive may not be in the child’s best interests with regard to reading the book themselves. Guidelines used by publishing companies to inform their font choices for children’s books often recommend the fonts most frequently in use and follow style trends rather than empirical data on legibility (Woods et al., 2005).

Magombe (2011) conducted interviews with representatives from major publishing houses regarding typography choices in children's books. The results (Table 3.5) indicate that decisions concerning fonts used for children are frequently made without the needs or reading ability of the child in mind. It is often assumed that the child will be read to by an adult, rather than developing readers attempting to read favourite books themselves. The publishers' comments align with the assertion of Sassoon (1993) that a book is often designed to appeal to the adults buying the book for a child rather than to help the child learn to read independently.

Table 3.5 Publishers' rationale for font choices in children's books (Magombe, 2011)

Publisher	Comment
Usborne books	Style and fashion are very important in typography, so we are continually trying new looks and typefaces. A lot of our decisions are made by gut instinct and by drawing on our experience as publishers.
Pearson Education	It is important that as a reader becomes independent, they are exposed to a variety of fonts so that they are prepared for the vast array of fonts that face them in all forms of media.
Walker Books	For our picture books the designers choose a typeface which they feel is the best one to complement the art and design of the book. They do not particularly worry about the child's reading ability because picture books are read to the child by an adult.

Phinney and Colabucci (2010) assert that typography should enhance the design of a children's book. They argue, for example, that if a story retells an ancient folk tale, it should not be set in a modern font, but one that evokes the impression of age. This raises the question of whether artistic impact is more important than legibility for young readers. A publisher may argue that a specific font that had been demonstrated to support developing readers may jar against the overall aesthetic.

This in turn raises the question of whether the design aesthetic of a book should override other considerations to the extent that a less ‘child-friendly’ font might be chosen.

Some publishers put legibility for the child before visual appeal, as publishers Heinemann use their own font for Early Years and Special Needs publications that has long ascenders and descenders and a modified ‘p’, ‘d’, ‘b’ and ‘q’ to help instant recognition (Magombe, 2011). The font was developed in conjunction with literacy advisors and Special Needs advisors. This indicates that this publisher at least is not just considering aesthetics and possible sales appeal, but also the clarity of the text to support children to read more independently.

One major difference in fonts intended for use in texts for children is the use of so called ‘infant’ versions of ‘a’ and ‘g’. The infant ‘a’ is designed to be like a handwritten version (ɑ) rather than the two storey version (a) commonly used in printed material. The ‘g’ is also like a written version (g) not a two storey version (g).

Teachers tend to believe that a sans serif font with infant characters is necessary for young readers. This belief has become embedded in tradition and is often followed by designers and publishers without further evaluation. Walker and Reynolds (2003) studied children’s reading performance when faced with the two versions of ‘a’ and ‘g’ and found that there was no beneficial or detrimental effect of the infant versions, which was later confirmed by Bessemans (2016) who states that two storey ‘a’s and ‘g’s do not create any added difficulty for children when reading.

Many typeface families have been extended to include a version that includes infant characters and are often delineated by the suffix schoolbook e.g. Century Schoolbook, AG Schoolbook and FF Schulbuch. Some fonts have been specifically designed for children e.g. Sassoon, Fabula and Fiendstar, which all have the infant style favoured by many for beginner readers.

Ripoll (2015) worked with a group of 120 five and six year olds in Spain and found participants read equally well with a serif or sans serif font, indicating that children may not need a font that is sans serif in all of their reading materials. Sassoon (1993) indicates that as children are exposed to a large number of fonts from an early age, especially if they read from screens as well as print on paper, then they subconsciously become aware of the many ways that letters can be formed. The publishers of children's books (Table 3.5) also state that a diversity of fonts is beneficial for a child to experience. Walker (2005) indicates that it is unlikely that every style of typeface will suit every child, so a variety of fonts should be introduced and should be a significant element in a child's book choice.

3.9 Specialist fonts and young readers with dyslexia.

The focus of this study is on the design of fonts and how this could support readers with dyslexia; hence it is beneficial to consider prior research on whether a specialist font that has been designed with such readers in mind has an effect on reading performance. Several studies focus on adult participants with dyslexia e.g. Appert and Truillet (2016), de Leeuw (2010) and Hillier (2006). Adults with dyslexia may have learnt to compensate for their difficulties and might be able to achieve reasonably fluent and accurate word reading but would still have phonological processing problems or other difficulties associated with dyslexia. This means that results quoted from these studies could be influenced by this learned adaptation and experience of different typography over the years. This might lead to different findings compared to those from those studies undertaken with young participants who may not yet have devised elaborate coping strategies.

Only a limited number of research studies have been undertaken that address the effect of font design directly on reading accuracy, reading rate or comprehension of young readers with dyslexia. These studies are summarised in Table 3.6. However, it is not possible to derive any overall conclusions as to whether a specialist font supports the reading performance of those for whom the fonts were designed.

Table 3.6 Review of research on the effect of specialist fonts on the reading performance of young readers with and without dyslexia

Authors	Language	Participants D=Dyslexia C=Control	Fonts	Dependent Variables	Methodology	Findings	Conclusions
Ossen 2012	Dutch	n=39 age range 9y-12y all with dyslexia	Dyslexie 12pt Arial 12pt	Accuracy Reading speed Comprehension Opinion	Same text, 2x week for 1 month at reading level of each participant. Same comprehension text and questions each session. Questionnaire for opinions	Accuracy; av 6.2 fewer errors Dyslexie Speed.Av 32 seconds faster Dyslexie Comprehension Av 2.3 fewer mistakes Dyslexie 75% stated preference for Dyslexie.	Same text used over time span. Repetition effect. No statistical analysis = no significance levels reported
Pijpker, 2013	Dutch	n=64, av. age 10y11m D high RL* 9, D low RL13. C high RL 30, C low RL 12 *reading level (RL)	High RL Dyslexie 9pt,Arial 10pt Low RL Dyslexie 10pt Arial 11pt	Accuracy: Errors made Reading Speed: Reading time in seconds	2 sets of 4 different meaningful texts (Dyslexie and Arial) for each level and group. Passage av. 400 words	Accuracy: *D High RL, significantly fewer deletion errors with Dyslexie *D Low RL, sig fewer errors all types with Dyslexie. *C high RL, sig fewer substitution errors. *C low RL, sig fewer substitution errors	Dyslexia group overall made significantly fewer errors using Dyslexie. Control group overall error reduction not significant, but substitution error reductions

						Speech: *D, high RL & low RL faster Dyslexie but not sig. *C high RL, faster Dyslexie but not sig. *C low RL, faster Arial but not sig.	were significant, Speed was not affected significantly by font
del Real Garcia, 2014	Spanish	n=10, Age range 8-10 All diagnosed dyslexia .	Arial Dixy (dyslexia-friendly) MeMinas (cursive)	Accuracy: Errors made Reading Speed: Words per 60 sec	40 different words & pseudo-words in each font. Listed across page. No coherent meaning to text.	Accuracy: Arial words av. 6 errors (err)/pseudowords (psw) av.12 err. Dixy words av. 3 err. psw av. 7err. MeMimas words av.7 err.psw av.10 err. Reading speed: Arial words av 49s, psw av,57s. Dixy words av.46s psw av. 56s MeMimas words av.51s psw av. 58s	Fewest mistakes using Dixy, significance levels not reported. Reading speed little variation
Zikl et al., 2015	Czech	n=150, D=75 av. age 10y 2m C=75 av. age 10y 1m	Arial OpenDyslexic(OD)	Accuracy: Errors made Reading Speed: Words read per minute	Imaginary language equivalent to Czech. Presented as sentences. Same text read in 2 fonts. Repeated 2 weeks later. Font order varied.	Accuracy: *D error rate marginally less using OD but not sig. *C error rate not influenced by font Reading speed: *D 1 word per min faster using OD, but not sig. *C reading rate constant.	Marginal benefit of OD for Dyslexia group but not significant. Effect thought to be greater on children with more serious reading difficulties. Little effect on reading speed.
Marinus et al., 2016	English (Australia)	n=39, Av. age 9y 8m Low progress readers	1)Dyslexie default spacing 14pt 2) Arial default spacing 16pt 3)Arial 1.5pt increased spacing 4)Arial1.3pt between words and 1.0pt within words increased spacing	Reading Speed: Words read per minute	4 texts of equivalent difficulty read in each condition.	*Dyslexie read sig faster than Arial when both default spacing Conditions1 & 2. *Dyslexie read sig faster than Arial Conditions 1 & 3 *Dyslexie and Arial no sig difference in reading speed Conditions 1 & 4	In default settings Dyslexie affects reading speed but under specific spacing conditions the benefit is removed.
Wery & Diliberto, 2016	English (USA)	n=12 Av. age 10y 11m. All diagnosed dyslexia	OpenDyslexic 10pt Times New Roman 12 pt Arial 12pt	Accuracy: Percentage correct Reading speed : 1 min reading time per list	Vertically arranged lists of letters, real words nonsense words	Accuracy: *No significant difference between OpenDyslexic or Arial and Times New Roman. *No significant difference in accuracy between times new Roman and Arial Reading speed; *No impact of font	No benefit reported for accuracy or speed for font design.
Bachmann &	Italian	n=533	Easy Reading (ER)	Accuracy: Errors made		a)Text	

Mengheri, 2018		Av. age 9y 6m. 3 groups 1) normal readers 426 2)reading difficulties 27 3)Dyslexia diagnosis 54 4)<25th percentile non-verbal IQ =26	Times New Roman	Reading speed: syllables per second	a) Meaningful text, b) list of words c) list of non-words all read in both fonts.	*Accuracy significant increase using ER for groups 2&3. *Reduction in accuracy for groups 1&4 using ER. *Speed increased for all groups using ER b)Word lists: *Accuracy sig increase using ER for groups 1&3. *Speed sig increase for groups 1&3. c)Non-word lists *Accuracy sig increase for groups 1&3, *Speed sig increase for groups 1&3	ER font has beneficial impact on accuracy and speed for children with dyslexia. Normal readers performance is improved with ER too. Spacing not included.
Duranovic et al., 2018	Bosnian	n=69, 23 D group Av, age 10y7m 23 Chronological Control (C) Av. age 10y4m 23 Reading Age (RA) Control Av. age 7y8m	1) Dyslexie default 11pt 2) Times New Roman (TNR) 14pt spacing increased by 1.3pt 3) TNR italic, 14 pt spacing 1.3pt 4) CurlzMT 14pt, spacing 1.3pt 5) TNR default 14pt	Accuracy: Errors made per syllable Reading Speed: Words per second	100 unrelated, meaningful short sentences, (20 per 5 conditions).	*D group read faster and with fewer errors using spaced text compared to default regardless of font design. *C and RA no impact on reading performance	Main impact found to be increased spacing not font design for D group. Bosnian is a transparent language so C and RA children not making reading errors by grade 1, age 6+, ceiling effect. Reading speed not affected for C and RA.
Kuster et al., 2018	Dutch	Exp1 n=170 Av age 9y11m All with dyslexia Exp2 n=147 D= 102 av. 10y4m C= 45 av. 9y2m	Exp1 Dyslexie 12pt Arial 13pt Exp2 Dyslexie 13 spacing 0.85 Arial 16pt sp. 1.15 TNR 16pt sp 1.15 Additionally:- Dyslexie 11 pt 1.0 vertical line spacing (vls) Arial 14pt vls1.30 TNR 14pt vls 1.30	Accuracy: Errors made Reading speed: time to read text Accuracy and speed: - number of words read correctly in 1 minute	Exp2 Short sentences, same for each condition and every child Exp 2 Card 1: list CV,VC & CVC words Card2: 4 & 5 letter words Card 3: multisyllabic words	*No impact of font on accuracy or speed. No order effect on accuracy or speed for font. Second reading faster regardless of font. *Number of words read correctly in 1 minute for D and C groups not impacted significantly by font design.	Reading performances of children with or without dyslexia are not better when read in Dyslexie font. Compare to Arial or TNR. Dyslexie does not affect reading negatively. Changing font type, size or spacing may benefit individuals
Powell & Trice, 2019	English (USA)	n=36 Av. age 10y3m All with dyslexia	Dyslexie 12pt default between line spacing 1.2 Times New Roman 14pt double spaced between line 1.7. Arial 14pt double space between line 1.7	Accuracy: Errors made Reading speed: time to read passage Comprehension: 3 questions	3 stories of 200 words at grade 4 (USA) level. Each story read in different font. Comprehension questions, recall, no inference	Accuracy: *No significant difference between fonts Reading speed: *No significant difference in time taken to read passages between fonts *Comprehension: No significant difference in	No effect of font. Spacing may influence findings but default spacing not compared. Non-standardised passages difficulty level not equal so influenced

						comprehension scores between fonts	results, stated by authors
Galliussi et al., 2020	Italian	n=128 Av. age 12.4 D=64 C=64	1)DF, specially designed font 14 pt, default spacing 2) DF with inter-letter spacing increased +0.98pt. and default interword spacing 3) DF default interletter spacing with interword spacing increased +3.78pt 4)DF interword and interletter spacing increased Standard font, based on Verdana 14pt default plus spacing conditions 2,3,4 as before	Reading accuracy: words read correctly ÷ number of errors Reading speed Syllables per second	8 different equivalent texts with 2x (font)x2 (inter-letter spacing)x2 (inter-word spacing) permutations	Accuracy: *no difference in reading performance for either group with DF or standard letterform. *no effect of inter-letter or inter-word spacing Reading Speed; *no effect of letterform. *Increased speed not found with increased spacing for D or C groups * For D group. speed reduced when DF presented with increased interletter spacing but no comparable increase in interword spacing	Data showed no positive effect of font design. No increase in accuracy or reading speed. Impairment of speed with DF font for dyslexia group under very specific circumstances
Joseph & Powell 2022	English	N=71 Age 8-12 D=37 C=34	Dyslexie 14 pt Calibri 20 pt	Reading speed: target word times and rapid automatic naming (RAN)	Naming letters RAN Reading appropriate level coherent passages including target words Eye tracking of eye movement and fixations	Dyslexie increased RAN for D and C groups. No effect of font on speed of word reading in passage	Dyslexie may benefit letter identification

Three of the studies, Ossen (2012), Pijpker (2013) and Bachmann and Mengheri (2018) found that a specialist font designed for those with dyslexia benefitted the reading performance of young people with dyslexia. Zikl et al. (2015) indicated that a specialist font may not benefit all young readers with traits of dyslexia, but only those with more serious reading difficulties. Conversely, Wery and Diliberto (2016) did not identify any benefit in reading with a specialist font in any of the conditions that they tested. Del Real Garcia (2014) indicated the specialist font investigated might be supporting reading performance but the study does not verify the findings fully with significant data. Additionally, Bachmann and Mengheri (2018) found that reading using a specialist dyslexia font also impacted positively on the reading performance of the typically developing young readers in the control group of their study. Joseph and Powell (2022) report that Dyslexie, as a font with distinctive letters, may benefit letter identification for both their control group and test group participants. Thus, studies show varied findings, which lead to contradictory conclusions.

As has already been outlined, spacing may affect the reading performance of young readers (Table 3.4). Default spacing conditions vary between different fonts and therefore spacing is closely related to font design. As the fonts designed specifically for readers with dyslexia tend to be more widely spaced, five of the studies included controls for spacing (Duranovic et al., 2018; Galliussi et al., 2020; Kuster et al., 2018; Marinus et al., 2016; Powell & Trice, 2019). Four of these studies assert that wider spacing has a beneficial impact on reading performance and it may be that the spacing was affecting the findings rather than the actual design of the fonts. Galliussi et al. (2020) identified a slightly different impact of spacing, as they found that if inter-letter spacing and inter-word spacing are not combined proportionately, then reading speed is impaired. This shows a negative role rather than a beneficial role resulting from spacing. All these research projects that include spacing have been reported since 2016, indicating that control of spacing alongside font design seems to have become more prominent in recent research.

The studies discussed were carried out using a variety of font choices, with different methodologies, varying sample sizes, and were also conducted in several different languages. The specialist fonts studied are Dyslexie (7 studies) and OpenDyslexic (2 studies), which are very similar in design plus one study for each of Easy Reading and Dixy. Additionally, Galliussi et al. (2020) developed D-F, a dyslexia-friendly font especially for their study and the font is not available commercially at present. Three of the studies do not declare the point sizes of the fonts used (Bachmann & Mengheri, 2018; del Real Garcia, 2014; Zikl et al., 2015) but those that do indicate a range in size from Dyslexie and OpenDyslexic both at 10pt to the largest Dyslexie 12pt. A majority of the studies use Arial and/or Times New Roman as control fonts, with Curlz MT being used additionally by Duranovic et al. 2018; MeMinas included by del Real Garcia (2014), and Calibri by Joseph and Powell (2022). In contrast, Galliussi et al. (2020) use a modified form of Verdana as a control font that was redesigned for their research. As already indicated in section 3.5b that with regard to planning research on fonts, Beier and Dyson (2014) and Herrmann (2011) assert that when fonts are compared the x-heights should be matched. Seven of the studies used different point sizes for the pairing of the specialist font and control fonts. However, the difference in font sizes is not always consistent e.g. Dyslexie 11pt and Times New Roman control font 14pt were compared by Duranovic et al., (2018), but Powell and Trice (2019) compared Dyslexie 12pt and control font Times New Roman 14pt. The inconsistency of font sizes etc. may indicate that although the findings may be valid for a particular study, the comparison between studies and any subsequent generalisation of results may be open to interpretation.

Methodologies vary considerably between the twelve studies considered. Participants may have been asked to read either passages of meaningful text, lists of words or nonsense words or combinations of these. It is possible that the research design may have influenced findings. Lists of words would require retrieval of phoneme-grapheme correspondences to decode, whereas coherent text would require that syntactic and semantic elements were also accessed. When reading passages were laid out conventionally the effect of crowding may have been felt, whereas a list of words might not cause flanking elements, other than neighbouring letters, to influence reading performance. Thus, findings concerning the effect of font design may be influenced by additional variables.

The level of difficulty of meaningful texts and how this aligns with the reading ability of the participants tends to vary. Whether this is matched to individual ability (Pijpker, 2013), cohort reading levels (Jospeh & Powell, 2022; Kuster et al., 2018; Marinus et al., 2016) or a single level for all participants (Bachmann & Mengheri, 2018; Galliussi et al., 2020; Ossen, 2012; Powell & Trice, 2019) may influence results, especially if a passage is challenging for some of the participants, thus affecting error levels and reading speed. The effect of repeated text reading was avoided in some studies by having different texts of equivalent difficulty for each data collection phase e.g. Marinus et al. (2016); Galliussi et al. (2020), but other studies controlled for repetition by having periods of time between phases of testing and re-reading the same text e.g. Kuster et al. (2018); Zikl et al. (2015). Ossen (2012) controlled for repetition by changing the order in which the 2 fonts were presented but did not alter the text content. In this case, it could be argued that repetition may have had a greater influence than font design in the reported overall reduction of reading errors and the increase in reading speed.

The language in which the studies were conducted should be regarded as another variable and hence comparison between the studies must take into account the differences caused by the orthography of various languages. Thus, the impact of font on reading performance may present differently in studies conducted in English, which has an opaque orthography compared to Italian, Spanish, Dutch, Czech and Bosnian which all have transparent orthographies. The impact of dyslexia on reading speed is a more noticeable factor when the language read by children with dyslexia has a transparent orthography, whereas an opaque orthography tends to lead to increased errors and reduced reading accuracy. This is emphasised by Duranovic et al. (2018), who state that the number of errors made by their participants was too small to show a significant difference between the fonts tested. The authors also clarify that they counted errors per syllable, rather than per word, but still found that the participants, especially those who were typically developing readers, made few mistakes (Duranovic, 2018). The transparent orthography of Bosnian means children tend to establish phoneme-grapheme correspondences soon after commencing reading instruction and consequently make few reading errors.

The sample size of participants with dyslexia in the studies also varies. They range from the largest group of 170 participants with dyslexia, who took part in experiment one of the study by Kuster et al. (2018) to a sample of 10 children, all with dyslexia, in the study by del Real Garcia (2014). This indicates that the findings from some studies may be more generalizable than others due to the size of the sample of participants. The participants with dyslexia were also identified in different ways and as there is no international, gold standard test or definition for identifying or diagnosing pupils with dyslexia, the criteria for inclusion probably varied between every study. Five of the studies did not include a control group of typically developing readers.

Consequently, it can be seen that although all of these studies are concerned with investigations that relate to the same topic, i.e. the impact of a specialist font for readers with dyslexia, all exhibit differences when the details of the research are examined and so the results do not give a definitive assessment of the effect of font design. Findings are no doubt valid for the conditions outlined in each study but are not necessarily generalizable to all situations.

3.10 Readers' preferences and opinions of fonts

A qualitative element in a research study provides a perspective that quantitative data cannot i.e. the human side of the study that includes information about readers' opinions and preferences. Children may become more invested in reading with a particular font if they feel that text is presented in a preferred or appealing font design. Motivation to read can be affected by typeface (Sassoon, 1993; Thiessen & Dyson, 2010), so it could be that if a young reader is consulted about the font design of reading materials then they might feel more inspired to read. Walker (2005, p.19) states "Children's views about typography are as valid as those of the teacher ". If a pupil were able to indicate a font preference for a reading intervention, then it may be possible to tailor support to provide maximum benefit.

Font design can also influence the perception of a reader's own abilities (Adee, 2012), with readers being reported to interpret instructions as being more difficult if presented in a hard-to-read font. This raises the question of whether a font design may plant the idea in young readers' minds that a book might be too difficult to read or understand if the text looked complicated because of the layout or font, regardless of the actual content. Bosman (2014) talks of the placebo effect of fonts on young readers, whereby a preferred font might give a child the idea that they can read better with that particular font, even if reading test scores show no initial difference. The belief that the font helps might influence motivation and make reading more pleasurable. Performance could then improve because of increased practice. King (2018) reports that pupils became more enthusiastic readers after they had read an edition of a book produced in a dyslexia-friendly font. This could have a knock-on effect regarding motivation to read once they realised that they could complete a whole book. Thus, interventions could be more effective if the reader believes that the font makes them a better reader.

Several studies using a number of fonts investigate if the font preference of young readers matches their reading performance. Reynolds and Walker (2003) worked with typically developing young readers who indicated an overall preference for Gill Sans rather than Century even though there was no significant difference in the number of reading errors made using either font. Bernard et al. (2002) compared reading performance and font preference of 27 participants aged 9 to 11, who read texts in Comic Sans Arial, Courier and Times New Roman. Results showed no significant difference in reading performance in terms of accuracy or speed, but a significant effect for typeface preference was found, with Comic Sans and Arial being perceived as easier and faster to read than Times New Roman or Courier. A preference for a larger point size of 14pt rather than 12 pt was also revealed. Thus, these extant studies indicate that preference is not driven by measured performance in these cases and children are not always accurate in their perception of their reading ability when reading using a variety of fonts.

Of the studies outlined in Table 3.6, four also reported the preferences of the participants involved. Kuster et al. (2018), who did not name the fonts to the children to avoid any unconscious bias, found that Arial was preferred to Dyslexie, but preference was not related to better reading in terms of accuracy i.e. the number of reading errors made. The results were computed using chi-squared tests, which confirmed that fewer participants than expected, both in the group with dyslexia and the control group, preferred Dyslexie. In the study by Zikl et al. (2015), some participants claimed to have the feeling that OpenDyslexic was easier to read but similarly some said that it made it less easy to read. The participants provided opinions subjectively and exact numbers for each preference are not quoted, but it is stated that the preferences did not align with impact on reading performance. Ossen (2012) reported that 75% of participants expressed a preference for Dyslexie font over Arial on their questionnaire. It is not stated in what way, if any, this result aligns with the 56.3% of participants, who made fewer errors overall or the 68.8%, who read more quickly. Conversely, Wery and Diliberto (2018) assert that none of their twelve participants preferred OpenDyslexic font. However, they do not report how this opinion was ascertained, which may have influenced the results. Kuster et al. (2018) indicate that the font choice of young readers may be influenced by the way that the font or the text is presented and the manner in which participants are asked about their preferences. Research carried out for the designers of Dyslexie font (attributed to University of Lille but with no indication of peer review) states that more than 70% of the participants preferred Dyslexie font to Arial. However, as this research does not give names of researchers and is published by the company promoting the use of the font and not in a journal, the results should perhaps be treated with caution.

Thiessen and Dyson (2010) assert that children may indicate a preference for a typeface or page layout based on what they think they should choose rather than their actual preference. Sometimes a reader who struggles may state a preference for a font that they think a fluent reader would choose. Instead of choosing a larger, more widely spaced typeface that is conventionally regarded as easier to read, they may feel that they should choose a smaller set type. Additionally, children may show sensitivity to type convention and not prefer a font because they think it looks 'babyish' (Thiessen & Dyson, 2010). Thus, perceived societal pressures may outweigh preference based on reading performance or perception of easier text presentation.

In the typographic and graphic design world preference is often assumed to link to aesthetics and the concept of fonts having different personalities, which are reflected in the content of the text (Heller & Anderson, 2017). The idea that fonts aimed at children should be, for example, 'friendly', 'warm' or 'smart' is often a factor influencing font choice (Çağlayan, 2009), but it is not confirmed whether these apparent personality attributes are perceived within the text by the children themselves or decided by the designers. Some fonts regarded as fun for children to read or as having 'happy' personalities are not always easy to read. Fonts, such as those associated with Disney may be hard to read, but because of the universal perception of the company being for children and families, the main font used is perceived as a happy, fun font with an emotional value that is more important than legibility (Gendelman, 2015). Consequently, if children are asked if they like a font they may agree because of its associations rather than the effect it may have on their reading performance.

Pastel (2011) contends that children are perceptive and interested in font design when consulted, so children’s preferences should be included when choosing a typeface for children’s texts. This could influence design and layout, as children may have opinions of typography which differ from those of the publishers or adults who make the choices of the fonts used in children’s books (Sassoon, 1993). Above all else, children require a font that is legible (Pattison, 2020), but they may have their own opinions about what makes a font legible, which should be considered.

3.11 Conclusions

This literature review indicates that the effect that font design, text layout and elements of legibility exert on the reading performance of young readers is complex. Many factors interact and individual readers react to font design in unique ways. Reading, a skill that many take for granted, is itself complex and multi-layered but fundamentally requires language in written form to be perceived and interpreted to extract meaning. It may be possible to support this process for children who are experiencing difficulties in learning to read by adopting a particular font that provides increased legibility and clarity. It can be seen that the literature does not provide a clear picture regarding the use of a particular font in supporting the reading of those with traits of dyslexia. The variety of methodologies, focus and language of origin do not lead to firm conclusions regarding the benefits or otherwise of a specialist font. The small number and limitations of the studies available indicate that more research is needed if a consensus regarding the optimal fonts to help support the progress of readers with dyslexia is to be identified. Conversely, if the benefits of a specifically designed font were to be refuted through further research then assistance could be focussed on interventions and the use of specialist fonts would not pose any distraction to structured and proven support.

Thus, the overarching aim of this study is to ascertain, by employing a sufficiently powered and well-balanced research design, whether the adoption of OpenDyslexic, an open source, dyslexia friendly font, impacts the reading performance of Key Stage 2 participants, a population with which I have most pedagogical experience.

This study investigates a comparison of

The impact of OpenDyslexic and Arial fonts on the reading performance of Key Stage 2 readers with dyslexia

by considering the following research questions.

RQ1: Does using OpenDyslexic font lead to higher reading test scores compared to Arial font with respect to 3 aspects of reading:
1a. reading accuracy,
1b. reading rate,
1c. reading comprehension?

RQ2: Does spacing influence test scores for reading accuracy,
reading rate and reading comprehension when texts are presented in the fonts OpenDyslexic and Arial?

RQ3: Does a preference expressed by participants for OpenDyslexic or Arial font align with higher reading test scores for reading accuracy, reading rate or reading comprehension?

Chapter 4: Methods

To address the research questions outlined at the end of Chapter 3, an embedded mixed methods design was used, in which qualitative data provide a supportive, secondary role to quantitative data. The primary quantitative data were used to analyse variables, formalise comparisons and investigate the reading accuracy, reading rates and reading comprehension test scores of the participants to focus on Research Questions 1a, b and c. Quantitative data were also used to address RQ2**.** Qualitative data, incorporating the views, opinions and feelings of the participants, which introduced a complementary perspective, were used to focus on answering RQ3. Both the quantitative and qualitative data were collected concurrently, with each participant completing the formal testing and then immediately providing their views by way of a semi-structured interview.

4.1 Mixed methods design

A rigorous, structured and planned use of quantitative and qualitative data to address the research questions was underpinned theoretically by pragmatism, a paradigm related to mixed methods, which considers the research question to be more important than the method or worldview that underlies the method (Tashakkori & Teddlie, 2003). Pragmatism rejects the extremes displayed in positivism and interpretivism and seeks the middle ground. Pragmatists recognise the physical world as well as the social and psychological world, with knowledge being constructed and based on the reality of the world we experience and live in (Robson 2011). Onwuegbuzie and Leech (2005) suggested that both quantitative and qualitative approaches have different strengths. Thus, the power of both methods was utilised to give a greater understanding of the research focus. A real world, practice orientated view drove the desire to find an answer to the research question rather than the study being driven by epistemological and ontological beliefs. The mixed methods design enabled the study to mirror ‘real life’. The quantitative data allowed statistics to show how reading performance varied when reading using the two fonts. The qualitative data allowed the opinions and reactions that the participants had towards the fonts to be included to give a broader picture of the impact not just on measured scores but also on perceived performance and preference, which could then impact on future reading test scores. All of the participants provided data for both the quantitative and qualitative phases. The two types of data were analysed separately and then findings from the qualitative data used to enhance the findings of the quantitative data. In this way, the three research questions were addressed to provide a holistic picture to consider the overarching aim of the study.

4.2 Participants

This study considers the impact of OpenDyslexic font on the reading performance of Key Stage 2 (KS2) readers with and without traits of dyslexia, as shown by scores obtained in tests of reading accuracy, reading rate and reading comprehension. The participants comprised two groups, the Dyslexia (D) group and the Control (C) group. Prior to the commencement of the main data collection, all participants completed the Single Word Reading Test (SWRT) (Foster, 2007) to indicate their current reading level. These scores were standardised to enable comparison and ensure that the correct reading test material was selected for each participant. An independent t-test was used to compare the standardised SWRT scores of the two groups to check that there was sufficient difference for them to be regarded as a test and control group. The SWRT scores for the Control group (M=103.29, SD=12.68) were significantly higher, t(76) = -6.97, *p*=<.001, than those of the Dyslexia group (M=80.85,SD=9.23).

Several participants were speakers of additional languages, (37% of total, D=12/40 and C=17/38), but all participants were confirmed by schools as having fluent English and had been at school in Britain for over two years. Talking to and working with the participants confirmed their competence in English. Language fluency was required to avoid any impact on reading accuracy, reading rate or reading comprehension due to oral language difficulties. This gave greater certainty that any errors made were not the product of incorrect pronunciation or misunderstanding due to underdeveloped skills in English.

All participants had normal or corrected to normal eyesight. Any participants who wore prescription spectacles to read were required to wear them during the data collection.

4.2a. The Dyslexia Group

This was a purposive sample of Key Stage 2 (KS2) pupils. All had been identified by the Special Needs Co-ordinators (SENCos) at their schools as having reading difficulties after an initial recognition of anomalies such as poor phonological processing, slow and inaccurate reading and spelling difficulties. All participants in the Dyslexia group were receiving targeted support for literacy difficulties. Although only 7.5% (3/40) had a formal diagnosis of dyslexia made by an educational psychologist, 87.5% (35/40) had been assessed using published screening tests that had identified dyslexic tendencies; the remaining 5% (2/40) had had concerns raised by staff within school that indicated a specific reading difficulty. They were receiving reading support but had not yet completed screening tests. Participants were recruited from 7 primary schools in London and the south-east of the UK, varying from a village primary school to a large urban, inner city school. The final sample in this group comprised 40 students (21 female, 19 male). The average age of children in this group was 9 years 10 months (range 7 years 6 months to 11 years 8 months).

Initially, it was considered that participants should be identified by administering a standard dyslexia screening test before being selected for inclusion in the study but certain schools expressed concerns that extra testing would put pressure on pupils and could undermine decisions made by the Special Educational Needs staff if findings contradicted those already forming a basis of support for a student. Therefore, I decided that all participants would be chosen based on being identified as needing support for dyslexia by their individual schools. The British Dyslexia Association definition of dyslexia (see Table 2.2) includes all of the difficulties that were initially recognised by class teachers and addressed through support interventions. This may mean that not all pupils had the same level of difficulties but teachers using their professional expertise had identified them all as having reading difficulties associated with dyslexia.

During the data collection, 2 participants in the Dyslexia group were identified by the SWRT as reading at a level commensurate with the York Assessment of Reading for Comprehension (YARC) Beginner level. The passages at this level are read partly by the teacher administering the test and partly by the child. The standardised scores for this level do not provide a reading rate score. The decision was taken to allow the participants to complete their sessions so as not to imply failure, but to exclude all of their results. Two extra participants were recruited who could read at level 1 and above, which provided a complete set of data for reading accuracy, reading rate and reading comprehension.

4.2b. The Control group

This group comprised KS2 pupils, who attended the same schools as their Dyslexia group counterparts, so could be termed a convenience sample. They were easy to invite to participate alongside the purposive sample participants and their inclusion allowed the distribution of request letters without drawing attention to the pupils with dyslexia. They were matched by chronological age to the Dyslexia Group participants and had age-expected level reading development with no identified reading difficulties as confirmed by class teachers and SENCos. This group comprised 38 participants (20 female, 18 male) with an average age of 9 years 6 months (range 7 years 7 months – 11 years 6 months). The Control group is two participants smaller than the Dyslexia group as one pupil was absent on the day of the data collection at their school, and one pupil decided against participating on the day. The decision of this pupil was respected in line with the BERA ethical guidelines (2018) adhered to during this study.

4.3 Ethics

Ethical considerations were of primary importance as the participants were children. Therefore, their status as a vulnerable group needed to be considered carefully (Robson, 2011). The study was registered with the UCL Data Protection Officer; number Z636410620190151, issued 11.1.2019, and the study was approved by the Institute of Education Ethics Committee on 22.1.2019.

The British Educational Research Association (BERA, 2018) requires that all participants should be treated fairly, sensitively, with dignity, and within an ethic of respect and freedom from prejudice, regardless of age or any other significant characteristic. The BERA (2018) ethical guidelines were adhered to during this study. The participants in any educational research should be allowed to give voluntary consent to take part in the project. Therefore, it is beholden on the researchers to ensure that the participants understand what they are being asked to do, why their participation is necessary and how the information will be used (BERA, 2018).

Participation was entirely voluntary with no inducements offered to encourage participation or penalties for non-participation. Written permission was obtained from each participant's parent or guardian confirming that they wished their child to be included in the group of children taking part[5]. The parents were given an information leaflet about the general nature of the study[6]. What was required from the children was explained in the permission letter.

[5] Appendix 2
[6] Appendix 3

Researchers must also respect the right of any participant to withdraw for any reason and at any time. The pupils themselves were asked if they wished to continue to participate once parental approval had been granted. Pupils received a verbal face-to-face explanation of what the study was about and what they were going to be asked to do if they agreed to participate. They also had a letter to give their own consent[7]. There was a smiley/unhappy face[8] provided to place on the table to indicate their decision or show if they wanted to withdraw at any time, after which their data would not be included. This method of communication was included in case the young participant felt too overwhelmed to state that they wanted to stop. If necessary they could just point to the sign or turn it over to indicate that they wished to withdraw. It is recognised that the use of this “stop” emoji may not have helpful connotations (Dockett et al., 2012). It was not the intention to imply that anyone would be unhappy with the child if they wanted to withdraw assent. It was explained to each participant that if the unhappy face reflected how they themselves were feeling at a particular moment then they could stop and that would have no repercussions.

Participants were reported anonymously, using pseudonyms that they chose themselves. Lahman et al. (2015) state that participants should be assigned human names rather than numeric identifiers. If a researcher allocates pseudonyms they might assign names that are attributed to certain characteristics or ethnicities, which may not reflect the participant, therefore the name should be chosen by the participant themselves (Lahman et al., 2015). Morrow (2008) asserts that it is conceivable that children who choose their own pseudonyms might pick a nickname

[7] Appendix 4
[8] Appendix 5

or name of a friend, which could make them identifiable to those who know them. However, in the case of this study, it is unlikely that a child could be identified by a friend’s name, as neither the names of the schools nor their exact locations are reported.

It was never the intention to imply in any way that the child’s real name was unacceptable in any way but pseudonyms were purely to protect identities. The children were asked to choose a pseudonym as a fun way for them to be both involved in the process, and to remain unrecognised. This was to protect the pupils and schools from any unintended consequences that could arise from identification. It was made clear that the use of a pseudonym did not imply the children were ever in danger by participating, but was part of an ethical process (Dockett et al., 2012). No documents included the participants’ real names.

Groups not individuals were reported on for the quantitative section of the study and quotes from individuals from the qualitative data were only attributed via pseudonym and not linked to schools or locations. The audio recordings were stored securely with password protection and were not shared with anyone. Raw data were also stored securely and only shared with my supervisors before analysis. Documents or data were not left in any of the schools or at university at any time.

The participants in the Control group were from the same schools and classes as the Dyslexia group so children could not be identified as being in a particular group when leaving the classroom to participate. Participants were not informed of the groups that they were assigned to for the data collection so that none were explicitly aware that they were being considered as having reading difficulties.

4.4 Researcher

I am a teacher with Qualified Teacher Status (QTS) and a clean enhanced DBS certificate, which was shown at each school. Other documentation was provided as required by individual schools. I familiarized myself with the specific safeguarding procedures of each school before the commencement of data collection and identified each safeguarding officer. I have never been employed by any of the participating schools, so the participants did not know me prior to their participation. In all schools the SENCo had told the children who I was and why I was coming to their school prior to my arrival, and in many cases introduced me personally to the children taking part. Although I was not the participants' teacher I could have been perceived by them as a teacher who was visiting and therefore in a position of power. The children were assured that they could withdraw without any consequences if they changed their minds about participating.

After receiving a request for their child to participate, it is possible that some parents could have felt compelled to consent, possibly because they see researchers or school as powerful and would not want to compromise their child by refusing. However, some parents decided not to give consent for their children to take part, so it was clear that they realised that refusal or withdrawal was possible and acceptable.

4.5 Data Collection

The data collection from all participants took place during the summer term of the same year. Each participant took part individually, so no other pupils were present to either distract them or influence their responses. Each session, one per child, lasted for approximately 30 minutes and took place in a variety of locations but all were outside the pupils' classrooms in small side rooms, offices, vacant classrooms or school libraries. All locations were as quiet as could be arranged with a table and two chairs side-by-side, and well-lit with a comfortable ambient temperature. None of the participants voiced any concern over the location that they were asked to work in.

To begin each session the study was outlined to each participant and the smiley face was placed on the desk to indicate that the child had given consent. The need for a pseudonym was explained and chosen by each participant. The reason for audio-recording of the session was clarified with permission to continue obtained verbally from each participant. A small Dictaphone was placed clearly on the table between the researcher and participant and used only for the data collection sessions and subsequent analysis.

During the data collection phase of the Institution-Focussed Study (IFS) (Broadbent, 2018), which acted as a pilot for this study, it was found that the name OpenDyslexic, which clearly implies the raison d'être of the font, might influence the participants' performance and perception. Consequently, the different fonts that the passages were reproduced in were referred to as "this font" and "that font" with the test examples being placed in front of the participant and pointed to. During the qualitative data collection, the sheets printed in the two fonts were laid out for the participants to see and again referred to as “this font” and “that font”. This ensured that the font names did not subconsciously influence the participants’ reactions to each font. This is in line with Kuster et al. (2018) who did not name the fonts that were used to their participants, mainly to avoid influencing a preference for any of the fonts that they presented. This prevented children from thinking that they should or should not prefer a font based on its name rather than its design or perceived benefits.

4.6 Typography and data collection

The style and appearance of the printed matter used in the data collection is fundamental to this study. The research questions focus on font design and spacing, as the overarching aim is to ascertain the impact of typography on reading performance. Thus the elements of reading accuracy, reading fluency and reading comprehension were evaluated by comparing test scores generated by each participant reading passages in two, specific fonts.

4.6a Choice of fonts

Every participant read coherent texts in both OpenDyslexic font and Arial font. Arial was chosen as a control font as it is a widely available, commonly used, sans serif font with good clarity and legibility. It was used previously as a control font in studies concerning the effect of font type by Del Real Garcia (2014), Kuster et al. (2018), Marinus et al. (2016), Ossen (2012), Pijpker (2013), Powell and Trice (2019), Wery and Diliberto (2016), and Zikl et al. (2015). Arial was argued to be an accessible font for people with dyslexia by Rello and Baeza-Yates (2013), which indicates that Arial is a suitable comparison to OpenDyslexic font. As argued by Beier and Oderkerk (2019) fonts that are compared should have all variables such as stroke width, x-height or weight controlled for. This was difficult to achieve with a font as individual as OpenDyslexic, in which the main design elements are the weighted bases and the uniqueness of every letter. Arial does not have the graded weight and some of the letter designs are reversals or mirror images of others. Thus, the two main factors of comparison are different. As outlined by Beier and Dyson (2014) and Herrmann (2011) when comparing fonts x-heights must be considered rather than point size. The overall heights from the top of the tallest character to the longest descender beneath the baseline of the letters in point size 16 of OpenDyslexic and Arial fonts are equal. However, the x-height, the distance between the baseline and the tops of the main body of the lower case letters, is marginally smaller in Arial font compared to OpenDyslexic font. Therefore, the point size used for Arial throughout this study was 16pt and the point size used throughout for OpenDyslexic was 15pt. The x-heights are nearer to being the same and the difference in overall height is almost imperceptible (see Table 4.1). The line spacing was set at 1.5 lines for both the texts printed in Arial and OpenDyslexic font.

Table 4.1 Comparison of Arial and OpenDyslexic fonts

Arial the quick brown fox jumps over the lazy dog	*16 pt*
OpenDyslexic the quick brown fox jumps over the lazy dog	*15 pt*

4.6.b Spacing

As indicated by the work of Duranovic et al. (2018), Galliussi et al. (2020), Kustler et al. (2018), Marinus et al. (2016), and Powell and Trice (2019), the impact of a specialist font designed to help the reading of those with dyslexia may be influenced by spacing rather than by font design per se. Before data collection began for this study each participant, whether from the Dyslexia group or the Control group, was randomly assigned to read Arial font in either the default condition or an expanded condition i.e. each participant read using either Arial default or Arial expanded, plus OpenDyslexic in the default condition only (Table 4.2). The aim was to determine whether any impact demonstrated between the two fonts was altered when the spacing was more comparable than when using both the default settings.

Table 4.2 Comparison of Arial font with expanded condition and OpenDyslexic default spacing condition.

Arial the quick brown fox jumps over the lazy dog	*16 pt default spacing*
Arial the quick brown fox jumps over the lazy dog	*16 pt, spacing expanded by 1.3 pt*
OpenDyslexic the quick brown fox jumps over the lazy dog	15 pt default spacing

The expanded Arial condition matched the spacing of the default OpenDyselxic font. This was achieved by the Arial font being altered using the ‘expanded’ function in Microsoft Word. Other more complex manipulations of intra-word and inter-word spacing, such as those used by Slattery et al. (2016) or Marinus et al. (2016) can only be achieved by specially developed, individually designed computer programmes. As I did not have the budget or expertise to achieve this, I decided to follow the lead of Perea and Gomez (2012) and to alter default spacing using the standard function available in Microsoft Word. The texts in Arial default and OpenDyslexic default were presented with line spacing of 1.5, as advised by Bohl and Hoult (2016). To make the comparison of OpenDyslexic default and the expanded control version of Arial as comparable as possible it was necessary to present the Arial Expanded texts with line spacing of 2.0 as shown in Table 4.2. Those who read Arial in the default condition comprised n=40; 20 Dyslexia participants, Control participants 20 (group name “Default”). Those who read using Arial expanded comprised n=38, 20 Dyslexia participants, 18 Control participants, (group name “Expanded”). Both the Default and Expanded group read OpenDyslexic in the default condition.

To ensure that any differences between the scores in reading accuracy, reading rate or reading comprehension were not influenced by the groups having differing reading abilities overall, the Default and Expanded groups were compared by running independent t tests and results which found no significant difference between the SWRT scores of those who read using default Arial (M=96.48, SD=14.53) $t(76) = 1.38$, $p = .171$, and those who read using expanded Arial (M=92.14, SD=13.52).

4.7 Quantitative data collection

To answer the research questions it was necessary to identify a research tool that would provide data concerning the elements of reading accuracy, reading rate to indicate reading fluency and reading comprehension. Lists of individual or nonsense words could demonstrate the impact of font on decoding or reading rate (Wery & Diliberto, 2016) but not on reading comprehension, so coherent meaningful texts were needed. The York Assessment of Reading for Comprehension (YARC; Snowling et al., 2009b) primary edition was judged to be a suitable tool for the quantitative data collection.

The YARC is a UK standardised test to assess accuracy of prose reading, reading fluency and reading comprehension, including literal and inferential meaning (Snowling et al., 2009b). These tests have been used for research purposes as well as assessments in primary schools (Snowling, 2013). Only two of the participating schools used the YARC and none of the participants had been tested previously at their current reading level using the YARC, so were not familiar with the texts used in the data collection phase. The Single Word Reading Test (SWRT; Foster, 2007) administered at the beginning of each session enabled the selection of an appropriate level of the YARC passages for testing each participant. This ensured that participants read a passage suited to their ability, which is not necessarily the same level as that expected for their school year. The participants read the printed passages aloud, which is a normal experience for primary school children and all were used to reading to a variety of adults such as class teachers, teaching assistants, volunteer reading helpers or family members. In this respect reading the passages aloud to me did not represent an unusual situation. Reading was scored for reading accuracy, reading rate and reading comprehension.

Each level of the YARC test has two parallel passages, form A and form B. The participants read both the required A and B passages from the identified levels, one passage in Arial font, either default or expanded condition depending on their allotted spacing group, and one in OpenDyslexic font.

Reading errors were noted using the guidance in the YARC manual. They were coded into categories as required by the YARC online analysis tool: i.e. mispronunciations, words that are wrongly pronounced or only partially decoded and have no meaning; substitutions, an incorrect real word given instead of the word in the text; refusals, when the reader is unable to attempt the word; additions, when the reader inserts a word or part of a word; omissions, when the reader omits a word; and reversals, when the reader reverses a word. The data produced was later analysed using paired sample t tests to ascertain whether the types of error recorded were made significantly more frequently depending on the font in which texts were presented. Spontaneous self-corrections were not counted as errors.

The time taken to read the passage was recorded to the nearest second. Immediately that the participant finished reading the passage, the eight comprehension questions were asked and answered orally. The second passage from the same level was read in the alternate font i.e. if passage 1 was read in Arial, then passage 2 was read in OpenDyslexic font. If the criteria outlined in the YARC manual were met, then the participant read passages from the following, more difficult level in both fonts and answered the comprehension questions. Two participants failed to meet the criteria to move to the following, harder level, so read passages from the preceding, easier level to the one that they had just read.

The levels alternate between fiction and non-fiction so each participant read fiction and non-fiction passages in both font conditions allocated to them. The length of the passages vary according to the reading level but range from 64 words on level 1 to 227 on level 6, with an average length of 167 words. In total each participant read four passages, two in each of the two fonts, and answered four sets of corresponding comprehension questions.

By using texts from the YARC, which have been trialled extensively, it is unlikely that there would be any variation in the difficulty of the passages of the same level that would influence any findings. Additionally, the passages and font conditions were alternated so that any previously unidentified discrepancy between apparently comparable texts would be controlled for. The first participant read using Arial font for the first passage and then OpenDyslexic for the next. Participant 2 began with OpenDyslexic, followed by Arial and then this pattern alternated to control for any order effect.

Data were prepared for analysis using the YARC online scoring conversion tool. Raw scores were computed to produce standard scores, which were then used to investigate the impact of the fonts.

Data were analysed using repeated measures Analysis of Variance (ANOVA), using the computer statistics package Statistical Package for the Social Sciences (SPSS) version 25. Repeated measures design was used because the same participants read using both Arial and OpenDyslexic font. A series of 2 (font type: Arial and OpenDyslexic) x 2 (group: Dyslexia and Control group) ANOVAs were computed to determine whether test scores in reading accuracy (RQ1a), reading rate (RQ1b) and reading comprehension (RQ1c) were impacted depending on the font used.

To investigate any influence of spacing, the data from the ‘Default’ spacing group and ‘Expanded’ spacing group were analysed using repeated measure 2 (font type) x 2 (default and expanded group) ANOVA for scores obtained in reading accuracy, reading rate and reading comprehension (RQ2).

4.8. Qualitative Data Collection

A short semi-structured interview, which had already been piloted in the IFS (Broadbent, 2018) was administered immediately after the completion of the YARC test to ascertain each participant’s opinions of the two fonts and their perception of their own performance[9]. This section of the total session of approximately 30 minutes lasted around 5 minutes. The interview questions allowed for participants to be guided towards providing answers that would generate useful data. Children may not be used to or capable of giving long, reasoned answers without structure (Walker & Reynolds, 2003) so a series of questions to provide a scaffold for their thoughts and opinions was asked to each participant. Nevertheless the interviewees were given space to define the answer in their own way to enable them to outline new and unexpected insights. The first question was intended to establish a rapport and to develop the initial relationship built up during the reading session. These answers were not intended for inclusion in the data analysis unless they revealed information pertinent to the research question.

Descriptive statistics were generated from answers obtained during the semi-structured interviews. They were produced to analyse expressed font preferences and to allow investigation of any alignment of preference with higher test scores

[9] see appendix 6 for questions used in semi-structured interview

achieved when reading using the preferred font.

Thematic analysis was used to analyse the qualitative data. This is a method for systematically identifying and organizing patterns of meaning across a data set (Braun & Clarke, 2012) and was employed because of its theoretical flexibility that would match a mixed methods design. It enabled the main focus to be on the research question rather than being driven by theoretical assumptions.

The comments and opinions of the participants were investigated to find patterns of meaning in relation to their reasons for their stated font preference. After the data collection the participants' answers were transcribed orthographically, reproducing all words and grammar. The childlike use of speech was retained. Non-word utterances such as "Hmm" were not included, as they did not bring additional understanding. The transcript was read analytically to ascertain how the participants made sense of the experience of reading using different fonts. The initial codes were outlined[10]. These codes had not been predetermined prior to the start of the study but developed during the reading of the responses. Words such as "letters" were identified as linking to different meanings, i.e. the shape or design of a letter, the clarity of the individual letters, the spacing of the letters and the impact of letters on reading performance. Hence, all of the codes were grouped into sub-themes dependent on meaning. These sub-themes were then combined to create coherent themes, which had a focus to explain how the relevant data answered the research question. The themes were developed from the participants' comments and gave an understanding of why the participants expressed their font preferences (see figure 5.2 in Results).

[10] Appendix 7 for example of analysis and coding of participants' comments

Thus quantitative and qualitative data were considered after analysis to compare performance scores in reading accuracy, reading rate and comprehension with participants' font preferences.

4.9 Summary

An embedded mixed methods design was used with the focus on answering the research questions. The participants were divided into two groups, the Dyslexia group all of whom were identified by their school as exhibiting traits of dyslexia, and the Control group, all confirmed as having age expected reading development with no identified reading difficulties.

The research tool was the York Assessment of Reading for Comprehension (YARC) and each participant read passages, matched to their reading ability and presented in fonts OpenDyslexic and Arial. Standardised, numeric test results for reading accuracy (RQ1a), reading rate (RQ1b) and reading comprehension (RQ1c) from both fonts were analysed to ascertain whether higher, overall scores were obtained when texts were presented in either font, by either group.

Impact of spacing on test scores was investigated by each participant reading texts using OpenDyslexic with only the default spacing and additionally either Arial default spacing or Arial expanded spacing (RQ2).

Each participant provided preferences and opinions of the fonts, which formed the basis of investigating RQ3.

Chapter 5: Results

This chapter reports the findings from the analysis of both the quantitative and qualitative data to address the research questions.

RQ1: Does using OpenDyslexic font lead to higher reading test scores compared to Arial font with respect to 3 aspects of reading:
1a. reading accuracy,
1b. reading rate,
1c. reading comprehension?

RQ2: Does spacing influence test scores for reading accuracy, reading rate and reading comprehension when texts are presented in the fonts OpenDyslexic and Arial?

RQ3: Does a preference expressed by participants for OpenDyslexic or Arial font align with higher reading test scores for reading accuracy, reading rate or reading comprehension?

5.1 Results: Quantitative data

Research questions RQ1a (reading accuracy), RQ1b (reading rate) and RQ1c (reading comprehension) were addressed by analysing the standardised data generated by the YARC online tool using repeated measures Analysis of Variance

(ANOVA), as the same participants read using both Arial and OpenDyslexic font. A series of 2 (font type: Arial and OpenDyslexic) x 2 (group: Dyslexia and Control group) ANOVAs were computed to determine whether there was a difference in test scores depending on the font in which the passage was presented.

5.1.a Reading Accuracy

Table 5.1 Descriptive statistics for reading accuracy

	OpenDyslexic		Arial	
	Mean	SD	Mean	SD
Dyslexia n=40	96.80	8.81	90.50	8.52
Control n=38	114.26	9.32	110.47	10.28
Total n=78	105.31	12.58	100.23	13.73

Table 5.1 presents the descriptive statistics for reading accuracy. A 2(font type: Arial and OpenDyslexic) x 2(group: Dyslexia and Control group) ANOVA revealed a significant main effect for font $F(1,76)=52.16$, $p<.001$, $\eta^2p=.407$ driven by greater accuracy when reading using OpenDyslexic font compared to Arial font. As interpreted using the published benchmarks for partial eta squared i.e. small effect 0.01; medium effect 0.06; large effect 0.14 (Draper, 2016; Richardson, 2011), the partial eta squared score indicated a large effect size. A significant effect of group was also observed $F(1,76)=90.01$, $p<.001$, $\eta^2p=.542$, indicating that the Control group read more accurately than the Dyslexia group. Finally, the interaction effect was not significant $F(1,76)=3.22$, $p=.076$, $\eta^2p=.041$, demonstrating that both groups were performing in a similar way, with both groups benefitting from higher test scores

in reading accuracy when reading in OpenDyslexic font.

Table 5.2 Comparison of error types made when reading using OpenDyslexic and Arial fonts

Type of error	group	mean	SD	t test
Mispronunciation	Dyslexia	OD 2.33 Ar 3.15	2.8 2.7	t(39) = -2.26, *p* = .014
	Control	OD 1.50 Ar 2.08	1.3 2.3	t(37) = -1.47, *p* = .149
Substitution	Dyslexia	OD 4.47 Ar 7.78	3.4 4.8	t(39) = - 4.88, *p*<.001
	Control	OD 2.53 Ar 3.61	1.9 2.8	t(37) = -2.58, *p* = .014
Refusal	Dyslexia	OD 0.88 Ar 0.70	3.4 2.5	t(39) = .293, *p* = .771
	Control	OD 0.00 Ar 0.00	0.0 0.0	Cannot be calculated as no control group participant recorded a refusal
Additions	Dyslexia	OD 1.00 Ar 1.38	1.5 1.8	t(39) = -1.56, *p*= .121
	Control	OD 0.87 Ar 1.26	1.4 1.2	t(37) = -1.81, *p*= .079
Omissions	Dyslexia	OD 1.45 Ar 2.38	1.9 3.2	t(39) = -1.96, *p*= .057
	Control	OD 1.26 Ar 1.71	1.5 1.9	t(37) = -1.51, *p*= .140
Reversals	Dyslexia	OD 0.05 Ar 0.20	0.2 0.4	t(39) = -2.23, *p*= .032
	Control	OD 0.03 Ar 0.00	0.2 0.0	t(37) = 1.00, *p*= .32

The types of reading errors were coded and are presented in Table 5.2. Within each group and within each error type, paired samples t-tests were used to test whether significantly different numbers of errors were made when reading text in the two different fonts.

Both groups made significantly fewer substitution errors when reading text in OpenDyslexic font. The Dyslexia group also made significantly fewer mispronunciation errors and reversals when text was presented in OpenDyslexic font. For the other types of error there was no significant difference between the two fonts for either group.

5.1b Reading Rate

Table 5.3 Descriptive statistics for reading rate

	OpenDyslexic		Arial	
	Mean	SD	Mean	SD
Dyslexia n=40	98.03	11.52	96.13	11.94
Control n=38	116.61	9.11	115.66	10.83
Total n=78	107.08	13.95	105.64	15.00

Table 5.3 presents the descriptive statistics for reading rate. The results of the repeated measures ANOVA test show a significant main effect for font $F(1,76)= 8.96$, $p=.004$, $\eta^2 p=.104$ indicating that the reading rate is faster when reading text in OpenDyslexic font compared to Arial font. The effect of the font on reading rate is not

as large as the effect seen on reading accuracy and is considered to be a medium effect as indicated by the benchmarks outlined for partial eta squared by Draper (2016) and Richardson (2011). There is a significant effect of group F(1,76)=61.55, p<.001, η^2p=.447 that shows that the Control group read significantly faster than the Dyslexia group. The interaction effect of group was not significant F(1,76)= 1.00, p=.320, η^2p=.013, thus both groups acted in similar ways, benefitting from higher test scores when the text was presented in OpenDyslexic font.

5.1c Reading Comprehension

Table 5.4 Descriptive statistics for reading comprehension

	OpenDyslexic		Arial	
	Mean	SD	Mean	SD
Dyslexia n=40	99.02	9.54	97.90	9.93
Control n=38	108.11	11.74	107.39	9.97
Total n=78	103.45	11.53	102.53	10.97

Table 5.4 presents the descriptive statistics for reading comprehension. Analysis using ANOVA confirmed no significant difference between the comprehension scores when reading using either font F(1,76)=0.78, p= .381, η^2p= .010 with a small effect size. There is a significant group effect F(1,76)= 19.74, p=<.001, η^2p= .206 with the Control group achieving higher scores than the Dyslexia group. The interaction effect was not significant F(1,76)=0.04, p= .843, η^2p=.001 so both groups were performing in a similar way with the comprehension scores of neither group benefitting from reading text in OpenDyslexic compared to Arial.

The main impact of font is that higher reading accuracy and reading rates were observed for both groups in the OpenDyslexic font condition. The effect of font was not significant for comprehension for either group. The Control group achieved a higher level in all three aspects of reading, i.e. reading accuracy, reading rate and reading comprehension compared, to the Dyslexia group.

5.2: RQ2: Impact of spacing

As outlined in section 4.6b, participants were randomly assigned to a group that read using either the default version of Arial (n=40, Dyslexia participants 20, Control participants 20, group name “Default”) or the expanded version of Arial, which closely matches the spacing of default OpenDyslexic, (n= 38, Dyslexia participants 20, Control participants 18, group name “Expanded”). All participants (n=78) read using the default version of OpenDyslexic regardless of allocated Arial spacing condition. Although the Control group participants in both the Default and Expanded groups scored more highly in reading tests than the Dyslexia group participants, when the Control group and Dyslexia group members were allocated to the Default and Expanded spacing condition groups, these two mixed-ability, spacing condition groups were shown to be similar in reading ability overall, as explained in 4.6b. The analysis comparing the scores generated when texts were read in OpenDyslexic and Arial indicated that there was no interaction between font and group when the groups were defined as Dyslexia and Control. Thus, it was possible to combine Dyslexia and Control group participants to create the Default and Expanded groups.

5.2a Spacing and Reading Accuracy

The reading accuracy results were compared to ascertain whether the increased spacing of the expanded Arial font influenced results. OpenDyslexic font had been found to have a significant effect on reading accuracy as outlined in section 5.1a and reading rate, section 5.1b. The aim of this analysis was to find whether the reading accuracy results were impacted by spacing by comparing reading accuracy scores of participants who read using OpenDyslexic and the default version of Arial and those who read using OpenDyslexic and the expanded version of Arial.

Table 5.5 Descriptive statistics for reading accuracy by font condition

	OpenDyslexic[1]		Arial[2]	
	Mean	SD	Mean	SD
Default group n=40	106.70	12.39	101.27	13.54
Expanded group n=38	103.84	12.76	99.13	14.02
Total n=78	105.31	12.57	100.23	13.72

[1] All participants read using OpenDyslexic default condition
[2] Participants read either using Arial default or Arial expanded condition depending on allocation

Table 5.5 presents the descriptive statistics for reading accuracy when comparing the Default and Expanded spacing condition groups. A 2(font type: Arial and OpenDyslexic) x 2 (Default group and Expanded group) repeated measures ANOVA was conducted. As expected the effect of font was significant, $F(1,76)=50.65$, $p<.001$, $\eta^2p=.400$, indicating that reading accuracy was higher in the OpenDyslexic font condition. However, there was no significant difference between Default and Expanded groups $F(1,76)=0.74$, $p=.392$, $\eta^2p=.010$. Nor was there a significant interaction between font and group F(1,76)= 0.25, *p*=.617, η^2p=.003. This indicates that the spacing condition of Arial had no reliable impact on reading accuracy scores.

5.2b Spacing and Reading Rate

The reading rate scores were analysed to find whether the Arial font spacing condition affected the reading speed of the participants.

Table 5.6 Descriptive statistics for reading rate by font condition

	OpenDyslexic[1]		Arial[2]	
	Mean	SD	Mean	SD
Default group n=40	106.78	14.48	104.83	15.60
Expanded group n=38	107.39	13.54	106.50	14.51
Total n=78	107.08	13.96	105.64	15.00

[1]All participants read using OpenDyslexic default condition
[2]Participants read either using Arial default or Arial expanded condition depending on allocation

Table 5.6 presents the descriptive statistics for reading rate and spacing conditions. This was analysed further using a 2(font type: Arial and OpenDyslexic) x 2 (Default and Expanded group) repeated measures ANOVA. As expected the difference in score between OpenDyslexic and Arial was significant F(1,76)= 8.97, *p*=.004, η^2p=.106, indicating that participants have higher reading rates when reading using OpenDyslexic compared to Arial. There was no significant difference between the Default group and the Expanded group F(1,76)=0.12, *p*=.726, η^2p=.002, indicating that the different spacing conditions of Arial did not influence the performance of the participants. No significant interaction between group and font was observed

F(1,76)= 1.24, *p*=.270, η^2p=.016. Overall, it can be seen that the spacing condition of Arial does not influence reading rate.

5.2c Spacing and Reading Comprehension

The initial investigation into the impact of font design on comprehension scores (see 5.1c) did not identify any significant differences in scores when using Arial or OpenDyslexic. It is logical to assume that neither Arial default nor Arial expanded would have a significant impact on comprehension test scores when compared to OpenDyslexic font. However, to check this assumption that comprehension scores are not impacted by font spacing, analysis was carried out in line with the other aspects of reading already investigated.

Table 5.7 Descriptive statistics for reading comprehension by font condition

	OpenDyslexic[1]		Arial[2]	
	Mean	SD	Mean	SD
Default group n=40	104.98	10.52	102.75	10.53
Expanded group n=38	101.84	12.46	102.29	11.56
Total n=78	103.45	11.54	102.53	10.98

[1]All participants read using OpenDyslexic default condition
[2]Participants read either using Arial default or Arial expanded condition depending on allocation

Table 5.7 presents the descriptive statistics for reading comprehension and Spacing conditions. Further investigation using ANOVA revealed no significant effect of font on comprehension scores, F(1,76)=0.74, *p*= .391, η^2p=.010; no significant difference between the Default and Expanded groups F(1,76)= 0.591, *p*= 0.444, η^2p=.008 and no significant interaction between font and group $F(1,76)= 1.68$, $p=.199$, $\eta^2p=.022$. Thus, spacing did not impact the comprehension scores of the participants.

In summary, spacing does not affect reading accuracy, reading rate or reading comprehension when comparing OpenDyslexic and Arial fonts.

5.3 Summary of quantitative study findings

The findings demonstrate that texts shown in OpenDyslexic font were read more accurately by participants in comparison to the Arial font. This was the case for both the Dyslexia group and Control group. The findings also show that texts presented in OpenDyslexic led to significantly higher scores for reading rate for the participants in the Dyslexia group and the Control group. However, font type did not significantly affect reading comprehension scores, thus participants did not score higher when reading in OpenDyslexic compared to Arial.

Spacing did not impact on reading performance terms reading accuracy, reading rate or reading comprehension when comparing OpenDyslexic and Arial in either the default or expanded condition.

These findings from the quantitative data will be discussed in Chapter 6 and possible reasons for the impact of OpenDyslexic on reading accuracy and reading rate. Its lack of impact on reading comprehension will also be considered. The lack of impact of the spacing manipulation will also be discussed.

5.4 Results: Qualitative Data

As outlined in Chapter 4 an embedded qualitative study supports the quantitative findings to allow the participants' voices to be heard. After the participants had finished reading their specific passages, they were shown the two fonts that they had been using side by side. The participants were asked to identify which font, if any, that they preferred. As the font names were not provided, most identified their choice by pointing and saying "that one", which was then recorded on the questionnaire schedule[11]. The participants were also asked their opinions of these two fonts.

35.4a Preferences

Table 5.8 Preferred fonts as expressed by all participants (n=78) by group

	Prefer Arial	Prefer OpenDyslexic	No Preference
Dyslexia n=40	11	23	6
Control n=38	17	16	5
Total n=78	28	39	11

Table 5.8 presents the font preferences of all of the participants. There is a clear difference between the two groups with respect to their preferred choice of font. In the Dyslexia group there was a noticeably higher number of participants, 12 more, who expressed a preference for OpenDyslexic rather than Arial compared to the Control group, where there was no clear preference for either font, with just a

[11] Appendix 7

difference of 1 more participant preferring Arial to OpenDyslexic. The participants who expressed no preference were also almost equal between the two groups.

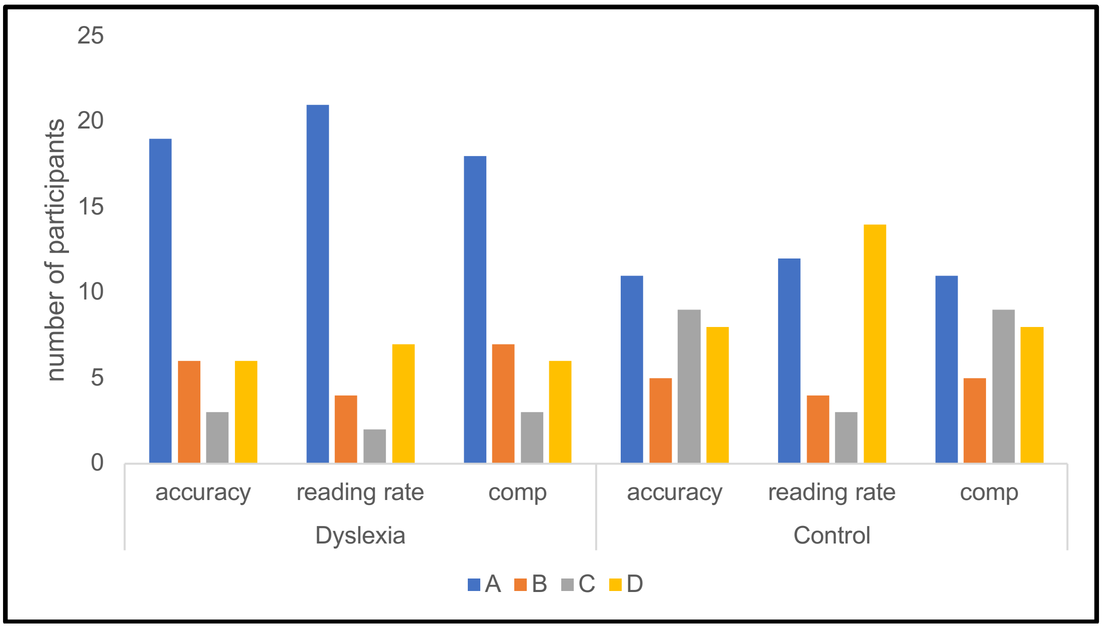
5.4b Preferred font and reading performance

To address Research Question 3 it was necessary to investigate whether participants preferred the font in which they achieved higher scores in terms of reading accuracy, reading rate or reading comprehension. A comparison was made between the choice of preferred font and reading performance as indicated by the results of the YARC tests.

There were 67/78 participants who stated a preference for either OpenDyslexic or Arial font; with 34/40 of the Dyslexia group and 33/38 of the Control group. 11 participants did not have a font preference, 6/40 of the Dyslexia group and 5/38 of the Control group.

Figure 5.1 shows the number of participants who preferred either OpenDyslexic or Arial font and whether their highest score in either reading accuracy, reading rate or reading comprehension was achieved in that preferred font or the alternative font. It also shows whether these participants were in the Dyslexia group or Control group. When considering the Dyslexia group participants whose preferred font and higher scores in reading accuracy, reading rate or reading comprehension aligned when reading using that preferred font, it is notable that more preferred and achieved higher scores when reading text in OpenDyslexic. However, there was also a preference /performance disconnect (see table 5.9) with 18 participants in theDyslexia group not scoring as highly in at least one of the measured test variables when reading text presented in their preferred font.

Figure 5.1 Alignment of Preferred Font and Higher Test Scores in Reading Accuracy, Reading Rate or Comprehension.



Legend

Dyslexia n=34 (n=40 minus 6 participants with no preference)
Control n=33 (n=38 minus 5 participants with no preference)

A= preferred font OpenDyslexic and higher test score using OpenDyslexic - preference and font align
B= preferred font OpenDyslexic but higher test score when using Arial - preference and font do not align
C= preferred font Arial and higher test score using Arial - preference and font align
D= preferred font Arial but higher test score when using OpenDyslexic - preference and font do not align

accuracy = reading accuracy
reading rate = reading rate
comp = reading comprehension

Figure 5.1 shows that among the Control group participants the alignment of performance and font preference was less pronounced especially regarding reading rate where a majority of participants read more quickly in the non-preferred font.

It appears from Fig 5.1 that for the Control group the alignment of font preference was the same for both reading accuracy and reading comprehension as the frequencies for categories A to D coincide. However, Table 5.9 indicates that the figures were generated by different participants in the two cases, so preference and reading accuracy and preference and reading comprehension were not aligning in the same way.

There were relatively few participants for whom preferred font aligned with improved scores in all three of the categories of reading accuracy, reading rate and reading comprehension (see table 5.9). From the Dyslexia group only 15/34 individuals scored higher when the text was presented in their preferred font in all three categories. Of these 15 participants, 14 preferred and performed better when the texts were presented in OpenDyslexic, 1 when in Arial. From the Control group 8/33 of the participants had higher scores in reading accuracy, reading rate and reading comprehension when reading using their preferred font with 7 preferring and scoring higher when reading with OpenDyslexic and 1 with Arial. Thus, it appears, especially for the Control group, but also for the Dyslexia group to a lesser degree, that font preference did not align overall with the scoring of higher test scores.

Table 5.9. *Alignment of Preferred Font and Higher Test Scores in Reading Accuracy, Reading Rate or Comprehension for each participant (Dyslexia group n= 34, Control =33)*

Preference for OpenDyslexic font					Preference for Arial font				
Participant[1]	group	accuracy	rate	comp	Participant[1]	group	accuracy	rate	comp
Laura	D	✓	✓	✓	Amelie	D	X	X	X
Alfie	D	✓	✓	✓	Natalie	D	X	X	X
Noah	D	✓	✓	✓	Emily	D	✓	X	✓
Jack	D	X	✓	X	Clark	D	X	✓	X
Lucky	D	✓	✓	X	Matilda	D	X	X	✓
Shayla	D	✓	✓	✓	Skyla	D	X	X	X
Daisy	D	✓	✓	✓	Unicorn	D	✓	✓	✓
Fred	D	X	X	✓	Dave	D	✓	X	X
Omar	D	✓	✓	X	Aidan	D	X	X	X
Chris	D	✓	✓	✓	Sophie	C	X	X	✓
Jamie	D	X	✓	✓	Sam	C	✓	✓	X
Lionel	D	✓	✓	X	Lily	C	✓	✓	✓
Erin	D	✓	✓	✓	Janet	C	✓	X	X
Melly	D	✓	X	✓	Steve	C	✓	X	X
Leah	D	✓	✓	✓	Amy	C	X	X	✓
Jeffrey	D	✓	✓	X	Alex	C	X	✓	X
Max	D	X	✓	X	Imogen	C	X	X	X
Charlie	D	X	X	X	Frankie	C	X	X	✓
Ben	D	✓	✓	✓	Scarlet	C	✓	X	X
Sophia	D	X	X	✓	Jake	C	✓	X	✓
Zoë	D	✓	✓	✓	Lauren	C	X	X	X
Jacob	D	✓	✓	✓	Lola	C	✓	X	✓
Oliver	D	✓	✓	✓	Jed	C	✓	X	✓
Shane	D	✓	✓	✓	Mike	C	✓	X	✓
Louisa	D	✓	✓	✓	Ramon	C	X	X	✓
Isabelle	C	✓	X	✓	King	C	X	X	X
Ava	C	✓	✓	✓					
Teddy	C	✓	X	X					
Messi	C	✓	✓	✓					
Rose	C	X	✓	X					
Tom	C	✓	✓	✓					
Tilly	C	✓	✓	X					
Emjay	C	✓	✓	✓					
Molly	C	X	✓	✓					
Saraiya	C	✓	✓	✓					
Peter	C	X	✓	✓					
Connor	C	✓	✓	X					
Mae	C	X	X	X					
Darcey	C	X	X	✓					
Jade	C	✓	✓	✓					
Lemon	C	✓	✓	✓					

[1]pseudonyms used for all participants ✓ = Font preference and higher score align X = Font preference and higher score do not align

5.4c Reasons for preference

The participants were asked to comment on the reasons for their preference for a particular font. These views were analysed to explore the reality of the experiences of the participants when reading using two different fonts. The comments were coded and then these were integrated to form themes.

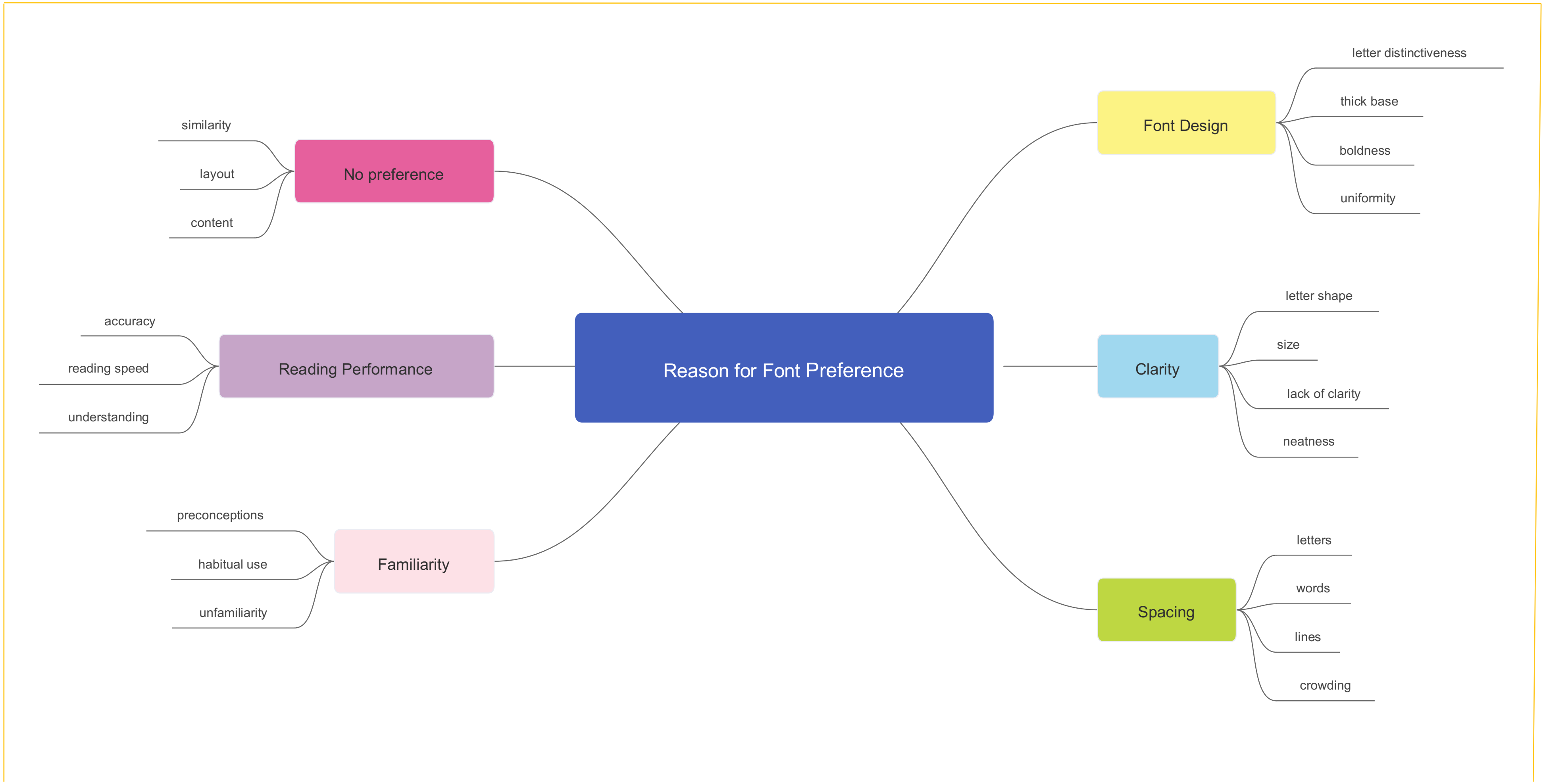
The data were analysed using an inductive approach as the themes were not outlined before the data collection and emerged from the data, although it could be argued that some themes were implied by the semi-structured interview. Questions were designed to steer the participants towards providing answers to the research question. The participants, being children, were largely too inexperienced to expand their thoughts on preferences without questions to guide and draw out their opinions. Thus, some themes emerged after being prompted by questions and others were arrived at more spontaneously and naturally. All themes link to the participants' perception of the benefit or detriment that a change of font had on their reading experience (see Figure 5.2). Themes are shown in coloured boxes in Figure 5.2 and sub-themes at the ends of the extending branches.

The six main themes outlined in Figure 5.2 provide the structure for reporting the findings of the factors that the participants indicated underpinned their preferences and opinions of the fonts.

When the qualitative data were analysed comments from certain participants appeared in several themes but other participants did not feature at all as they had made minimal contributions to this aspect of the study. Many of the participants spoke in broken, ungrammatical sentences, reflecting their thought processes as they considered their answers. This is demonstrated in the verbatim comments quoted.

The responses that the participants gave when questioned about their preferred font varied in length and depth of analysis. Some were extremely engaged in the comparison of the fonts and dissected the differences between the fonts and explained their decisions in detail. Conversely, some chose a preferred font but were unable to expand on precisely why they had made that decision. When sub-themes are discussed the number of participants who alluded to that category will be outlined, but for brevity only one illustrative quote will be given rather than listing all of the relevant quotes.

Figure 5.2 Thematic map illustrating themes and sub-themes that emerged concerning reasons for font preference



i)Font Design

The most frequently mentioned attributes that underpinned font preference were aspects of font design. The sub-themes are shown in Fig 5.2.

Each letter in OpenDyslexic is designed to have a unique shape, with none being the reversal or inversion of another letter, i.e. b and d and b and p are different shapes.

This can be regarded as letter distinctiveness. Although this feature was not conveyed to the participants, it was noted and commented on by some participants,

> e.g.EmJay (C)[12]: *The letters [OD][13] are easy to read because they are all different to each other.*

Ramon noted that the letters were not designed as might be expected but his interpretation could indicate that the letter shapes are not a positive factor.

> Ramon (C): *The stick of the b doesn't go down. It looks like they aren't finished but the d stick does.*

Jeffrey, who explained his preference whilst indicating that he habitually confused not just letters but also numbers, linked the distinctiveness of the letters to the amelioration of difficulties with letter identification.

> Jeffrey (D): *The coloured in at the bottom makes obvious letters [OD]. I can tell the O is not a zero. I can tell the d and b are not the wrong way round.*

Many of the participants cite the thick base of the OpenDyslexic design as a factor in their preference. Nine of the Dyslexia group regarded it as a positive attribute e.g.

---

[12] (C)= control group (D)= Dyslexia group

[13] Participant indicated they were referring to OpenDyslexic = [OD] or Arial = [A]

Leah (D): *The letters have more of a black line at the bottom. That helps. I like that.*

Conversely, seven of the Dyslexia group regarded the thick base as a negative feature. These participants dislike an element that makes OpenDyslexic unique e.g.

Skyla (D): *I like the shape of the letters but not the thin to thick part. If it was all thin it would be all right.*

The Control group participants were also divided in their opinions of the thick base. Four regarded it as a benefit e.g.

Peter (C): *I like the thick bits at the bottom of the letters. You can see what you are doing better.*

But six disliked the thickened base, e.g. Imogen, who felt it led to confusion rather than clarity.

Imogen (C): *I don't like the thick bottom then thin top. The change confuses me and you can't see what the words are supposed to be*.

Jake expressed similar confusion with the varying stroke thickness but focussed on Arial.

Jake (C): *Arial is clearer and all the same thickness. It means you don't confuse the letters.*

Several terms were linked to the variation in thickness of the OpenDyslexic letters as participants indicated that they felt the font stood out using a variety of words, e.g. "inky", "black", "dark", e.g.

Oliver (D): *It [OD] is more inky so I can see the words easier. It's much clearer.*

The control group tended to use the word "bold" to describe OpenDyslexic font e.g.

Tony(C): *The bold stood out more so it emphasises what you are reading.*

"Bold" was not always used to attribute a positive element and could be the cause of dislike of the font design e.g.

Sam (C): *I don't like those dark bold bits at the bottom. I like [A] a lot more.*

Not all of the participants who explained their preference did so in terms of liking or disliking an aspect of OpenDyslexic font. Some spoke positively of the design of Arial, mentioning the uniformity of the letter shapes. This approach was adopted mainly, but not exclusively by the Control group participants e.g.

Scarlet (C): *When letters are all the same like that [A] it is much better.*

Conversely, only one of those from the Dyslexia group mentioned uniformity as the reason for their choice.

Aidan (D): *That [A] is much more all the same. I think the word is uniform. That one [OD] the letters aren't. Some look really weak.*

ii) Clarity

When talking about font preference, participants mentioned frequently that their preferred font appeared clearer. Some of the participants inspected the text pages in detail and gave opinions on clarity based on individual letter shapes, thus linking clarity and design e.g.

> Tom (C): *The letters look clearer [OD]. I think it was designed to be clearer to read.*

Clarity was also linked to print size in some cases, even though the point sizes were manipulated so as to appear as equal as possible e.g.

> Zoë (D): *It's clearer. It looks a bigger size [OD].*

Whereas most of the participants indicated a preference due to increased clarity, some indicated that a lack of clarity in one of the fonts lead to a preference for the other e.g.

> Lauren (C): *That [OD] is shadowy and unclear but that [A] is much clearer.*

Three of the Control group defined their preference in terms of neatness, suggesting that they assessed the overall effect of the font on the entire page. They all preferred Arial e.g.

> Steve (C): *That [A] is set out much neater. It is straight up. That [OD] is not neat because the letters are all more curled up.*

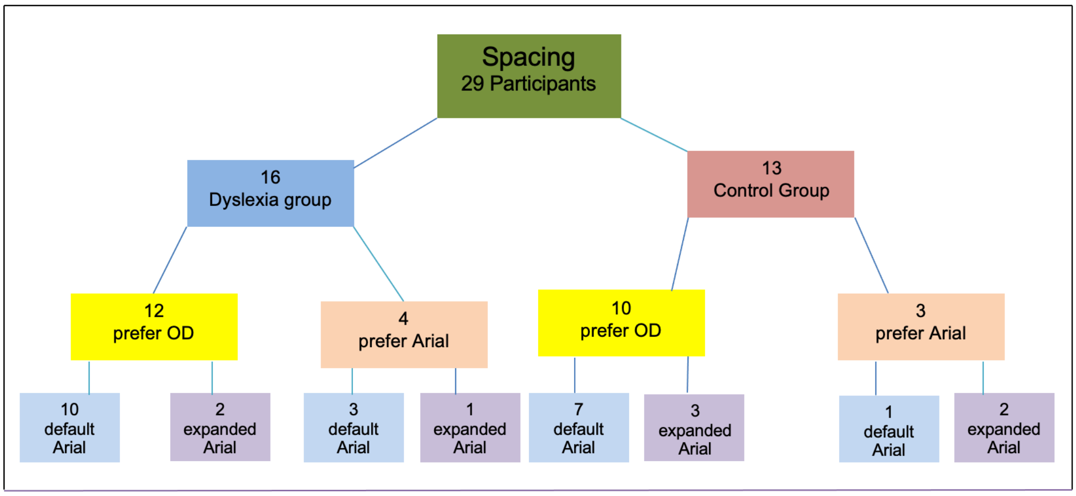
iii) Spacing

29 participants (16 Dyslexia group, 13 control group) defined their reason for their preferred font in terms of spacing. It was also noted which Arial font condition that

the participants had been exposed to, to ascertain if the spacing of Arial had influenced their decisions.

The number of participants, who cited spacing as a reason for their font preference, was similar between the Dyslexia group and Control group (see Fig 5.3). The participants from both groups preferred OpenDyslexic to Arial by a ratio of approximately 3:1

Figure 5.3 Breakdown of participants who cited spacing as a reason for font choice and the Arial spacing condition that they were assigned



The majority of those who preferred OpenDyslexic had read with Arial in the default condition, which could suggest that the more compact default spacing of Arial was influencing their decision in a negative way as well as the expanded default spacing of OpenDyslexic impacting positively. This is reflected in some of the comments. Some focussed on the layout of words, e.g.

Messi (C): *There are more spaces and gaps [OD]. I like that.*

Others considered the spacing of the letters, e.g.

Alfie (D): *That font [A] is too squashed up with letters to read properly*.

The comments of three of the Dyslexia group revealed that they may be experiencing the phenomenon of visual crowding, whereby recognition of letters is impaired by the presence of other neighbouring letters and textual features. These participants all referred to words running together and struggling to identify the boundaries of words e.g.

Lionel (D): *For me two words often go together but with this [OD] I can see spaces, so I know if it is one word or two by each other, not one together.*

iv) Familiarity

Familiarity of font design was cited by eleven participants (4 Dyslexia group, 7 Control group) as a reason for their choice. Ten of these participants preferred Arial. None of the participants who commented on familiarity had knowingly seen OpenDyslexic before. Four participants preferred Arial, as it was a font that they were used to, e.g.

Emily (D): *I prefer [A]. I am used to what it looks like.*

Conversely, three stated that they did not like OpenDyslexic as they were not familiar with it, e.g.

Sophie (C): *I have never read it before. I don't like it but I would have to see it more to decide.*

A further three participants preferred Arial as they associated it positively with text content, indicating preconceptions of how print should be presented, e.g.

Skyla (D) *I would choose it [A] because it usually makes a good story.*

The perception of how print should look was also linked to a layout and design that they were accustomed to, e.g.

Frankie (C): *All the writing in our school books looks like that [A] so I am used to it*

One participant preferred OpenDyslexic because of its unfamiliarity. This participant approached the new font with a wish to embrace the unfamiliar and clearly had a mindset to enjoy something new.

Jade(C): *I like the look of [OD] because I see [A] every day and I like to experience things that are different.*

v) Influence on performance

Most participants who considered their preferred font in terms of its impact on their reading performance used general phraseology such as “made it easier” which was used by 21 participants (13 Dyslexia group, 8 Control group). Only 8 participants made more specific comments on how the fonts impacted their reading. Three participants linked font to reading accuracy, mentioning how OpenDyslexic helped, e.g.

Tilly (C): *The letters are easier if you need to sound out the letters in new long words.*

Or that Arial was problematic, e.g.

Saraiya (C): *With [A] the letters are close so you can make mistakes*.

Only one participant from the Dyslexia group considered reading speed as a factor in his choice of font, e.g.

Chris (D): *When I read this one [OD] I could read quicker and not make mistakes.*

In contrast, 5 of the Control group mentioned reading speed as an influence on their choice. Two preferred OpenDyslexic e.g.

Isabelle(C): *[OD] makes the lines stand out and I think you can go faster.*

This is in comparison to three participants who preferred Arial e.g.

King(C): *[A] gets you reading quicker.*

Comprehension was also included in aspects of reading performance that font design could benefit e.g.

Molly(C): *Maybe if it [OD] makes me slow down I can understand better.*

Jacob linked his font choice to being able to read with fewer regressions. He indicated that this made his comprehension improve.

Jacob (D): *[OD] I don't look back as much. I can understand it better because I don't need to keep going back.*

vi) No Preference

The eleven participants (6 Dyslexia group and 5 Control group) who stated that they had no preferred font should not be overlooked as they all justified their reason for not having a preference.

A perceived similarity of the two fonts was cited by 4 participants (1 Dyslexia group and 3 Control group) as a reason for not stating a preference, e.g.

> Ocean (D): *I like both of them. I can cope with both of them. I don't really see much change.*

Text layout was mentioned as a factor in the decision to not express a preference by 3 participants (2 Dyslexia group, 1 Control group) e.g.

> Angel (D): *I like what I read to be broken up into paragraphs. The fonts are both ok for me.*

Additionally, 4 participants (3 Dyslexia group and 1 Control group) indicated that the reason for preferring the two fonts equally was that the content of the text was more important than font design, e.g.

> Rock (C): *I have no preference. I just want a book to be challenging. Both fonts are perfect for reading.*

It can be seen that many reasons were given for font preference with several factors influencing the preferences that were not connected directly to reading performance but were linked to typographical design elements.

5.4d Summary of findings of qualitative element of study

86% (67/78) of participants expressed a font preference with 58% (39/67) of those participants preferring OpenDyslexic. Many reasons for font preference were given, with the most frequently mentioned alluding to font design or spacing. A preferred font does not align with improved reading performance i.e. a higher score in reading accuracy, reading rate or reading comprehension is not consistently achieved when the text is presented in the preferred font. Thus, pupils chose a preferred font for a variety of reasons not necessarily linked to reading achievement.

5.5 Summary of findings

RQ1: Does using OpenDyslexic font lead to higher reading test scores compared to Arial font with respect to 3 aspects of reading

RQ1a. reading accuracy

A significant effect of font on accuracy was found for both the Dyslexia group and Control group when reading using OpenDyslexic font. Reading accuracy test scores were found to be significantly higher when the texts were presented in OpenDyslexic compared to Arial font.

RQ1b. reading rate

A significant effect of font on reading rate was found for both groups when reading using OpenDyslexic. Participants scored significantly higher when the test texts were presented in OpenDyslexic compared to Arial.

RQ1c. comprehension

Font was found to have no significant effect on reading comprehension test scores for either group.

RQ2: Does spacing influence test scores for reading accuracy, reading rate and reading comprehension when texts are presented in the fonts OpenDyslexic and Arial?

There was no significant difference between scores achieved in any aspect of reading by the group who read using Arial default condition or those who read with the expanded Arial condition. The benefit of OpenDyslexic compared to Arial was the same for both Arial conditions. Thus, it was not spacing that conferred an advantage when reading using OpenDyslexic font.

RQ3: Does a preference expressed by participants for OpenDyslexic or Arial font align with higher reading test scores for reading accuracy, reading rate or reading comprehension?

Participants did not necessarily prefer the font in which the text was presented when they achieved a higher score in reading accuracy, reading rate or reading comprehension tests. An expressed font preference was often not linked to improved reading performance in terms of test scores attained when reading using that preferred font. A variety of reasons were given for font preference that included factors other than the influence of the chosen font on reading performance.

Chapter 6: Discussion and Conclusions

The current study was conducted to ascertain whether a particular font could impact the reading performance of Key Stage 2 readers, with particular attention on pupils who exhibit traits of dyslexia, the rationale being that if a font were to be suggested to help young readers in schools then that recommendation would be based on quantifiable and empirical research. The study was underpinned by three research questions. The findings generated by the quantitative data which answered research questions 1 a,b,c and 2 are considered within section 6.1. The embedded qualitative aspect of the study which addressed research question 3 is discussed in section 6.2. Strengths and weaknesses of methodology are then outlined plus the role of the study in shaping professional practice.

6.1 Quantitative Strand

6.1a. Reading Accuracy

The findings of this study show that when text was presented in OpenDyslexic font reading accuracy scores were significantly higher. The large effect size demonstrated that the relationship between the variables of font type and reading accuracy scores was strong enough for the results to be considered meaningful. This benefit of font was true for both participants with dyslexia and the control group of typically developing readers, who had not exhibited any reading difficulties. This finding indicates that not only does font design impact reading accuracy for children with dyslexia but also impacts children, who whilst not having reading difficulties are still developing their skills as readers.

The findings that the reading accuracy of both groups is impacted beneficially support the findings of a study by Bachmann and Mengheri (2018). The participants of their study had a similar average age of around nine years six months and their study also used meaningful text presented in a dyslexia-friendly font and a comparison standard font. However, there are several differences. The fonts used are different, being dyslexia-friendly EasyReading and Times New Roman (Bachmann & Mengheri, 2018) in contrast to OpenDyslexic and Arial (this study). Visually EasyReading is very different to OpenDyslexic. However, the premise behind both of these fonts is that the design increases legibility. Letters that in some fonts are reversals of each other e.g. ‘p’ and ‘q’, ‘b’ and ‘d’, are designed in both EasyReading and OpenDyslexic with unique shapes so that they can be distinguished easily. The fact that EasyReading and OpenDyslexic look different from each other but both aim at increasing legibility may indicate that making a text more legible by using a font with unique, clear, easily distinguished letter shapes influences reading performance with regard to accuracy. This also indicates that in the dyslexia-friendly fonts the critical features that need to be identified to distinguish particular letters may be more easily detected. Thus, the text is more legible, leading to fewer reading errors.

As already discussed in section 3.4, letters are theorised as being identified as a series of features, not as a whole. It is possible that the dyslexia-friendly fonts have a clarity that enables the elements that are crucial to identification to be distinguished more easily and rapidly, even for relatively competent but still progressing readers, as well as those experiencing difficulties. The fact that readers with dyslexia and typically developing readers benefited from the dyslexia-friendly fonts also indicates that children who are still developing and improving their reading skills regardless of whether this development is at an expected rate can be supported by text being presented in a font with increased legibility.

The type of errors that the participants made revealed that both the Dyslexia and Control groups made significantly fewer substitution errors when reading text presented in OpenDyslexic font compared to Arial. The most common type of reading error was a substitution error, with the target word substituted with an incorrect but real word. The Dyslexia group also made significantly fewer mispronunciation errors, i.e. read fewer real words as nonsense words with no actual meaning; and made fewer reversals when the text was read in OpenDyslexic font. The Control group made few mispronunciation errors overall, so a significant difference between the fonts could have been masked by very small numbers.

In this current study findings regarding the type of error are comparable in general terms with the findings of Pijpker (2013). Participants in that study had an average age of 10.5 years and read meaningful, reading level appropriate texts in Dutch. The dyslexia group was sub-divided into children with a low level of reading ability (n=13) and a higher level of reading ability (n=9). A control group with no diagnoses of dyslexia were also divided into lower (n=12) and higher reading (n=30) ability groups. Regarding the two groups overall, there was a significant positive impact on reading accuracy when the text was presented in Dyslexie compared to Arial for the dyslexia group but not the control group. However, if the sub-groups and the type of errors are considered then the impact of font was more subtle. The participants who were identified as having dyslexia and low reading level made significantly fewer errors in general with most error types identified being impacted when reading text in font Dyslexie compared to Arial. Those in the group with dyslexia who had been identified as having a higher reading level made significantly fewer consonant deletion errors specifically when reading text presented in Dyslexie, i.e. missed out fewer consonants from words. There was not a significant impact of font on other types of error. Both of the control sub-groups, regardless of reading level made significantly fewer substitution errors when reading texts presented in Dyslexie font, but other types of error were not significantly impacted. Thus, Dyslexie font was impacting the number of errors, but more specifically on the type of error depending on the level of reading ability of the participant. In the current study, OpenDyslexic did not impact all types of reading error in the same way (table 5.2) with both groups making significantly fewer substitution errors when reading text in OpenDyslexic font compared to Arial, and the Dyslexia group also making significantly fewer mispronunciation errors and reversals, when text was presented in OpenDyslexic font.

Conversely, other studies outlined in Table 3.6 and with a variety of test fonts, carried out with a variety of methodologies, and in a variety of languages reported no significant impact of font design on reading accuracy (Galliussi et al., 2020; Kuster et al., 2018; Powell & Trice, 2019; Wery & Diliberto, 2016; Zikl et al., 2015). This could be due to a number of factors. In languages with a transparent orthography when the grapheme-phoneme correspondences are consistent e.g.Czech (Zikl et al. 2015), Dutch (Kuster et al., 2018) and Italian (Galliussi et al., 2020) readers with dyslexia tend to make fewer decoding errors than readers with dyslexia whose language has a more opaque orthography. This means that difficulties associated with dyslexia are less likely to manifest as poor reading accuracy. This indicates that a comparison of reading accuracy between fonts is not possible as too few errors are made for a comparison to demonstrate a significant difference. English has an opaque orthography whereby dyslexia can present as poor reading accuracy. In studies in English reading accuracy differences between fonts may be masked by methodological decisions such as using word lists rather than meaningful texts (Wery & Diliberto, 2016). Decoding of individual words does not generally include semantics or the need to make cohesive links with other words, so reading accuracy may not be impacted by the text content in the same way as when meaningful texts are used. The reading level of the passages used may also impact reading accuracy if they are not matched to the participants' reading ability e.g. Powell and Trice (2019), who used the same passages for all participants. If the passages presented are too easy or too difficult the comparison of reading accuracy between fonts may be compromised. An easy passage would reveal few errors regardless of font and a difficult text would have a considerable level of errors in all font presentations. It is difficult to compare the findings on reading accuracy of this study with extant studies as variables are not consistent. It would be necessary to replicate methodologies for clear comparison.

6.1b. Reading Rate

The findings show that there was a statistically significantly faster reading rate when the test texts were presented in OpenDyslexic font compared to Arial. The effect size is smaller than when testing font against reading accuracy but is still medium to large (Draper, 2016), which indicates a meaningful relationship. The beneficial impact of OpenDyslexic was again present for both the Dyslexia and Control groups. The

Control group read significantly faster than the Dyslexia group when reading text in either font, but nevertheless read significantly more quickly when using OpenDyslexic compared to Arial.

These findings are again consistent with those of Bachmann and Mengheri (2018), who demonstrated that the font EasyReading helped participants with and without literacy difficulties to achieve a significantly faster reading rate. Sousa (2014) asserts that reading pace gives an idea of the effort required to read, so if a passage is read less slowly and less laboriously then this indicates that the reader requires less cognitive effort to decode. Reading quickly could again be due to a clearer, more legible font helping with visual processing. This precipitates faster identification of letters and grapheme - phoneme matching, or allows words in the mental lexicon to be accessed more quickly, which consequently leads to quicker decoding because of improved legibility. Ossen (2012) also reported that participants read more quickly using the dyslexia-friendly font, Dyslexie. However, it is not reported whether these results are statistically significant so it is difficult to judge the robustness of the results. Joseph and Powell (2022) found that letter naming appeared more fluent when the letters were presented in Dyslexie rather than Calibri, indicating that the clearer letterforms aided letter identification even though they did not observe a knock-on effect of increased word reading speed.

Contrary to these findings, other studies outlined in Table 3.6 found no significant difference in reading rate scores when comparing fonts (del Real Garcia, 2014; Duranovic et al., 2018; Galiussi et al., 2020; Joseph and Powell, 2022; Kuster et al., 2018; Marinus et al., 2016; Pijpker, 2013; Powell & Trice, 2019; Wery & Diliberto, 2016; Zikl et al., 2015). All of these studies have many different variables, such as font choice, methodology and a number of participants, but none identified significantly different reading speeds when comparing their chosen fonts. It should also be noted that reading rate does not equate completely with reading fluency, as outlined in section 2.5b, so an impact on automaticity or prosody may not be identified within the data of studies looking purely at reading speed.

Research has shown how reading speed may be influenced by the size and frequency of saccades, length of fixation, number of regressions and any eye movements atypical for reading (Franzen et al., 2021). Font design may influence the layout of the text and this may encourage readers, especially those with traits of dyslexia, to process words as whole units by using fewer saccades and fixations as indicated by the work of De Luca et al. (2002). This could lead to an increased reading rate. However, this hypothesis cannot be tested without eye-tracking investigations, which is outside the remit of this study.

6.1c. Reading Comprehension

The effect of font design was not observed for reading comprehension in this study. This could be because comprehension does not begin until after letters have been perceived and decoding starts. At least some graphemes need to have been decoded before the reader can begin to access the words and then the meaning. Reading comprehension requires more regions of the brain to be active than when decoding without comprehension is happening (Coch, 2010) because the process of reading comprehension is more complex than decoding. Reading comprehension calls on additional skills such as visualisation to construct a mental image of the text; working memory and the executive function to maintain and update information as the text is processed; and reasoning and inference. To enable comprehension to be achieved the reader needs the cognitive capacity to access vocabulary, cultural knowledge and prior experience. As more complex processes are taking place in reading comprehension, font design is probably less likely to impact at the same level that it does decoding. There are more factors at work than the perception of letters and their mapping onto graphemes to decode words. Thus the role of the legibility of the font is less influential as it is competing with many other processes that lead to understanding a text.

The findings support those of Powell and Trice (2019), who also found no difference in reading comprehension scores when the text was presented to young readers in different fonts. However, Franzen et al. (2019) considered the role of OpenDyslexic font on the reading comprehension of adults. Contrary to the findings of studies with child participants, they found that there was a beneficial impact on the test scores of both participants with and without dyslexia. If the font design impacts legibility, which in turn allows easier decoding, then more cognitive capacity is available that can be used to comprehend and thus comprehension scores increase. To juxtapose the findings of Franzen et al. (2019) with the current study it could be argued that those adults were already able to decode with accuracy and automaticity, so more cognitive capacity became available for comprehension when font design had the effect of making text more legible. In the case of young, developing readers, or young readers with dyslexia, the extra cognitive capacity freed by the increased legibility of the font may be used to allow effortful decoding to become more accurate or automatic and thus comprehension is not impacted. The amount of cognitive load freed when the text is presented in OpenDyslexic compared to Arial may be sufficient to support more accurate reading and faster reading rate for both the Dyslexia and Control groups but is not sufficient to impact on comprehension.

It could be postulated that a less legible font would demand more attention to enable decoding and, thus, comprehension would be impacted adversely, as increased cognitive capacity would be needed to identify letters. However, there is a school of thought that implies that a dysfluent or 'hard to read' font enables deeper processing and consequently greater recall indicating improved comprehension (Diemand-Yauman, 2011; French et al. 2013). It is unlikely that OpenDyslexic is acting as a dysfluent font, even if it were initially unknown to readers, as its design aim is to be clear and legible. Hence, it is unlikely to impact comprehension in the manner suggested that occurs with 'hard-to-read' fonts and it would not impact comprehension test scores of young readers. The studies of Diemand-Yauman et al. (2011) and French et al. (2013) are concerned with comprehension demonstrated by recall of what had been read, whereas the passages used for the comprehension element in this study included questions that required inference. This category of question is less likely to be impacted by font as many cognitive functions work in conjunction to answer inference test questions. If OpenDyslexic were to act as a dysfluent font this effect may be masked by the type of comprehension questions being asked.

6.1d Spacing

Extant studies indicate that it is possible that a dyslexia-friendly font supports children's reading accuracy and reading rate by having greater spacing between the letters than a standard comparison font such as Arial. To test this hypothesis passages were presented in OpenDyslexic and Arial fonts in the default settings to 40 of the participants, the Default group (section 4.6b) and test scores were compared with the remaining 38 participants, comprising the 'Expanded" group, who read texts presented in default OpenDyslexic and the expanded Arial condition (Table 4.2). Findings showed no significant difference in reading accuracy, reading rate, or reading comprehension scores between these two groups. This indicates that the increased spacing of Arial font was not outweighing the benefits of OpenDyslexic default spacing. If spacing were playing a significant role, it would be expected that the level of reading accuracy and reading rate scores would be equal when the text was presented in Arial expanded, as the assumed benefit of OpenDyslexic being created by the wider default spacing would have been negated by the wider Arial spacing. However, findings show that this was not the case. The fact that findings for Research Questions 1a and b show that reading accuracy and reading rate when text was presented in OpenDyslexic default were significantly higher than when presented in either condition of Arial font, for both the Dyslexia and Control groups, indicates that the benefit is afforded by variables other than spacing.

This finding is not consistent with any of the other studies considered so far. All published research that identifies the benefit of a dyslexia-friendly font reports that once spacing was controlled for then the identified benefit was no longer valid (Marinus et al., 2016; Powell & Trice, 2019). Additionally, Kuster et al. (2018)

reported that no benefit of a specialist font, namely Dyslexie was found, regardless of default or increased spacing being presented. Conversely, Galliussi et al. (2020) found that adjusted spacing, rather than acting beneficially, impaired reading speed if the balance of inter-letter and inter-word spacing was upset. Bachmann and Mengheri (2018) did not include a control for spacing in their published study, but state that in further studies this would be done and speculate that controlled for spacing would impact results because of the crowding effect experienced by many readers with dyslexia. This is an assumption that they have not yet tested and so is not confirmed.

With regard to studies that did not include a specific dyslexia-friendly font, not all report that wider spacing was beneficial. Van den Boer and Hakvoort (2015) and Reynolds and Walker (2004) report that there was no significant benefit when spacing was wider than the default. This was the case for both typically developing readers and readers with dyslexia. Perea et al. (2012) and Zorzi et al. (2012a) assert that wider spacing benefited readers with dyslexia in terms of improved accuracy and faster reading speed but did not impact the performance of their control groups of typically developing readers. This diversity of findings indicates that there is no consensus regarding spacing and reading performance.

The fact that findings regarding spacing in this study do not coincide with findings of extant studies could be due to the way spacing was manipulated, the overall methodology of the study, or the characteristics of the sample of participants. In this study, the spacing was altered in Microsoft Word using the standard function for expanding text spacing. Other studies with different results manipulated spacing using MATLAB (Marinus et al., 2016) or a specifically designed, individualised computer programme (Slattery et al., 2016). It could be that in this study manipulating spacing using Word was not nuanced enough to preserve all of the additional variables inherent in the overall font design and so the subtleties of the impact of spacing were lost.

Variations in the design and methodology of studies may make comparison of the findings less robust. When considering the layout of text e.g. presenting a full, coherent text, individual sentences, or lists of words, these methodological choices may influence the role of spacing, especially with regard to the role of crowding. The findings of studies with a design using vertical lists (e.g. Wery & Diliberto, 2016) may not yield findings that are comparable with studies where text is presented with several words on a line and several lines in a passage. Words in a coherent text or when lists are printed horizontally tend to be closer together than in vertical lists, even if spacing is manipulated. Thus, the effect of peripheral letters or words impinging on perception may be reduced by layout rather than the impact of manipulated font spacing.

The point size of the font may also impact spacing. Most of the studies considered here used a smaller point size font than this study e.g. Marinus et al. (2016) used Dyslexie 12 point and Arial 13 point, or Powell and Trice (2019), who used Dyslexie12 point and Times New Roman 14 point, compared with this study wherein the text was presented with OpenDyslexic 15 point and Arial 16 point. In typographic design spacing and x-height are linked by ratios designed to maximise legibility. Small fonts need wider spacing. Thus, it could be that the large point size used in this study is less susceptible to impact by spacing manipulation than a smaller point font which needs a wider spacing to be legible. The smaller font may be more responsive to the impact of the alteration of spacing.

The characteristics of the participants may also impact the results of spacing. The studies considered in Table 3.6 might possibly include a substantial proportion of readers among their participants with traits of dyslexia, who are susceptible to the effects of crowding. If this were the case, then due to the characteristics of those participants, spacing may be playing a greater role in impacting the results than if a smaller proportion of participants were affected by the effect of crowding. The size of the sample also impacts findings, as in a small sample each individual has a larger role in shaping the results and greater influence on statistical findings. If the results of both readers with dyslexia and any typically developing readers in a control group are analysed together, the impact of those reacting to spacing and crowding may be further diluted and as such skew findings. Thus conflicting results from various studies may be found that are valid for the participants of that study but do not allow direct comparison.

Powell and Trice (2019) indicate that a font includes not just letter shape but also other characteristics including default spacing when it is trademarked, patented or copyrighted. Thus it could be argued that spacing is part of the font only in the designed, default setting. If spacing is altered then the totality of the font design is being corrupted. If a font is more legible than another in its default setting, it should be regarded as more legible overall. It is arguably not the same font if a particular variable is altered. Thus, researchers should not claim that the benefit of a font is removed if spacing is altered because the font is then a different font from how it was conceived and designed. By having a default that is widely spaced, the needs of the reader are supported by that font in its totality not just by the wider spacing.

Legibility is not just the shape of the letters. It includes the full package of elements that make up the font. Whether the efficacy of OpenDyslexic is due to letter shapes or its spacing, it is a complete package that may benefit the reader. There is an amalgam of variables that is impacting the reader. If it is acceptable to isolate spacing to explain findings, then all other intrinsic or extrinsic variables (Walker et al., 2018) should also be isolated and analysed. It is possible to argue that a font with wider default spacing may support reading performance. Spacing should not be divorced from design by implying that spacing is the only factor to benefit reading. This would indicate that all fonts are equally legible but variance in legibility is only due to spacing.

Beier and Oderkirk (2019) assert that fonts can only be compared if all variables, apart from that being tested, are identical. This might indicate that to test for the effect of spacing the same font must be used in different spacing settings, rather than comparing different fonts with apparently similar space settings. Thus, to test the impact of spacing in this study an additional aspect where participants read with texts presented in default OpenDyslexic and a condensed version of OpenDyslexic to counteract the wide default spacing would possibly have been more appropriate. As it stands, the effect of spacing could not be isolated to show impact on the reading performance of the Dyslexia group or the Control group. It should be concluded that the impact on test scores by OpenDyslexic font was due to the font design in its entirety and not just spacing.

6.2 Qualitative Strand

The qualitative element of this mixed methods study allows the perspectives and interpretations of the participants to provide greater insight and understanding of the role of font. Given the ultimate aim of the study is to underpin practice and support advice given to schools that impacts the pupils, the opinion of the children in the study should be included. If it transpired that the children themselves did not want to read using a particular font then recommending it purely on increased reading scores would be counterproductive in the long term. Thus the embedded design enables the opinions of the participants regarding their reading experience when using each font to be incorporated into the study. It allows each one to be valued for their individual contribution and not just playing a role as a generator of statistics. The participants expressed views that are respected and validated by inclusion in the study (Walker, 2005). Preference is personal to each participant.

With regard to the descriptive statistics it appears that the Dyslexia group had a strong preference for OpenDyslexic font (68%, 23/34) and the Control group did not have such a clear preference (48%,16/33). Overall, the preference for OpenDyslexic was 58% v Arial 42%. Even though OpenDyslexic font is designed primarily for readers with dyslexia it is justifiable to present texts in OpenDyslexic to readers who do not have dyslexia as shown by stated preferences. If OpenDyslexic were only suitable for readers with dyslexia, then it could be assumed that none of the Control group would prefer it as it would not appeal to them in a way that it might to the

Dyslexia group. However, the preferences indicate that if texts were to be presented in OpenDyslexic they would also be acceptable to readers without identified reading difficulties. When considering participants' preferences, from which the assumption can be made that they would be willing to read using that preferred font, the bigger picture shows that OpenDyslexic is suitable for all.

Participants did not necessarily consider the impact of either font on their reading performance when choosing a preferred font but gave a variety of reasons for their choice. The overriding factors tended to be linked to what was actually on the page i.e. the design, clarity and spacing of the font, and typography. This could be because contemplating how the font affected performance is a step removed from what is seen on the page i.e. the child might say for example 'I like that font, it has clear letters' but not include the next step of a possible thought process and say something along the lines of 'so I think the clearer letters make me read better when using that font'. Thus, it could have been the case that further detailed questioning could have teased out these nuanced thoughts regarding reading performance. However, findings would still have shown that the initial preference was not necessarily linked to better test scores. The non-alignment of preference of font to reading performance was also reported by Zikl et al. (2015) and Kuster et al. (2018), which indicates that young readers do not or are not able to link font design to how it impacts their reading.

Thiessen and Dyson (2010) assert that children will often express an opinion that they think is required or expected of them. It is possible that some of the participants expressed the preference that they thought most desired for the study. The fonts had not been named to the participants but it is quite possible that they could work out which font was which. Many of them would recognise Arial as a font they had seen before even if they did not know it by name and then deduce that the other font was the one that the study was focussed on. Consequently, the fact that the fonts were not named could have been overridden by prior knowledge and raises the possibility that some of the participants were trying to be as helpful as possible and provide the answers they thought were required. Conversely, it could have been possible that some participants chose Arial as their preference as they did not want the association of choosing a dyslexia-friendly font. Thus a further undeclared factor maybe operating in the background to influence font preference.

Emotions may have influenced the preference if the participant had worked out which font was the dyslexia-friendly font. Before the participants gave their own consent to be included, the purpose and methodology of the study was explained to them and they were aware of the underlying reasons for the research. Consequently, they may have thought along the lines of ‘I have dyslexia so that font helps me because it is specially designed for me’ or ‘I do not have dyslexia so that font won’t help me’ or ‘I do not want the dyslexia font to help me because I don't want people to know I have dyslexia’. All of these imagined scenarios could lead to a subconscious or unarticulated reason for preference that not declaring the font names explicitly had failed to prevent.

The choice of font preference is unlikely to have been influenced by peer pressure as the children who took part in the data gathering sessions did so on a one-to-one basis. No other children were present to influence a choice or exert pressure.

It must also be considered whether the participants were able to fully articulate their ideas, which raises the question of whether they had the vocabulary to actually express and explain what they felt. The intention of words such as ‘clearer’ is on the surface obvious but does not expand on what element of font design actually makes the font clear. This would be extremely difficult for many adults without graphic design knowledge to explain, so the young participants would find this even more difficult. It is also possible that a participant had a preference but did not know the reason why they had that preference, as it may have manifested as just a feeling. Some of the children did express a preference but were reticent to explain why. The reason could have been, not that they did not want to engage but that they did not have the conceptual and linguistic tools to do so.

The children were not asked for a preference until they had read the text presented in both fonts, so it is unlikely that preference had impacted reading scores. Although they had not voiced a preference at the point of reading the texts, they may have had a subconscious preference once they had seen both fonts. If participants had been shown both fonts beforehand or warned that they would be asked about their preferences, this may have influenced findings, as they may have had this in their minds for longer, rather than it being an immediate decision. When the participants were asked their preference they had already finished reading so any conscious or subconscious impact on reading performance would already have been experienced.

OpenDyslexic font was designed with improved legibility at its core. It has letters with unique shapes, clear negative spaces and well-proportioned ascenders and descenders. Each of the features outlined in section 1.3 was alluded to by at least one participant in this study, indicating that they are able to perceive the particular elements aimed at supporting readers with dyslexia through font design. Thus, the important features of the font are impacting the choice of preferred font. It should not be forgotten that some children perceived these unique identifying details as a negative, but nevertheless they were still perceptive to the design details of the less preferred OpenDyslexic font.

6.3 Strengths and weaknesses of the study

The methodology of the study could play a role in influencing the findings. By considering aspects of methodology this possibility can be examined and strengths and weaknesses outlined.

6.3a Participants

The participants could only be recruited from schools willing to take part. The schools were recruited via SENCos, who were all interested in engaging with research. This means that the schools had an ethos whereby they were open to new ideas and innovation to support their pupils. It may be that this mindset has influenced the pupils so they are more proactive and positive, possibly giving more effort when asked to participate as either they or their peers have already had exposure to the concept of investigative insider research. Schools that do not habitually engage with research, either by participating or keeping abreast of new innovations, are unlikely to want to be involved. This means that participants tend to have a predisposition towards a positive mindset regarding involvement so this may impact through enthusiasm and commitment to the procedures.

The parents or carers who gave initial permission for participation may also have had certain characteristics which might impact the sample of participants. They all agreed for their child to participate and thus saw participation as a positive opportunity. This might be reflected in the way that the participants engaged with the activities, bringing with them the influence of adults who regarded the opportunity in a positive light. Some parents who were approached did not consent, so their children did not take part, which may have influenced the sample. Some types of research that include entire class cohorts as part of a learning objective may include participants that have more negative reactions towards research participation. This may impact outcomes in subtle ways that if multiplied by large numbers of participants affect findings.

6.3b.Groups

Whether the participants were included in the Dyslexia group or the Control group depended largely on the information provided by individual schools as indicated in section 4.2. The participants in the Dyslexia group were not tested specifically for inclusion in this study before being selected but most had been tested in school before being recruited, with all receiving targeted reading support in school, which was the qualifying criterion for inclusion. As the participants were not identified by one particular test but selected by their schools, they might have had a wider range of difficulties than had they all been tested using the same testing regime. However, analysis of the scores generated by the Single Word Reading Test (SWRT), administered prior to data collection, indicated that overall the two groups were significantly different, with the scores for the Control group being significantly higher than the Dyslexia group, thus demonstrating that the reading level of the Control group was higher than that of the Dyslexia group. All of the participants in the Control group scored above 85 and so had a standard SWRT score within the average range. All but three of the dyslexia group scored below 85, outside the average range. The participants that scored within the average range appeared to do so by adopting individual strategies for decoding single words but still exhibited traits of dyslexia when reading coherent text. The groups were sufficiently dissimilar for a comparison of findings to be valid. The fact that both groups were acting in similar ways regarding effect of font on reading accuracy scores and reading rates indicates that the font was impacting the findings rather than the reading level of the groups being too similar to allow results to be distinguishable.

6.3c. Sample size

The size of the sample was 78 participants in total, Dyslexia group n=40, Control group n=38. This approximates the recommendation of 40 participants per group (Budiu & Moran, 2021) with the Control group having 2 participants withdraw at a late stage. This allowed for a confidence level of 95% for computing a margin of error. This sample size was large enough to allow parametric statistics to be used, unlike the IFS (Broadbent, 2018) study which had 12 participants in total and in which data were analysed using non-parametric statistical methods. The extant studies considered in Table 3.6 have a median sample size of n=64, which is comparable to this study.

The sample size was determined by the limited resources available, namely the number of schools willing to take part; the number of potential participants within those schools, especially those fitting the criteria for the Dyslexia group; and the number of participant permissions agreed. Time was also an influencing factor, as it was only possible to work with a maximum of 6 participants during a standard school day, depending on the timing of lunch and breaks plus other non-negotiable times when the participants had to be with their cohort. Some schools restricted the amount of time that they were willing to allow pupils to be released for participation e.g. not during core subject lessons. Additionally only one researcher was collecting all of the data so this dictated time restrictions and no funds were available for travel. Consequently, limited resources had to be choreographed to deliver the largest feasible and practical sample. A larger sample may have produced different or more representative findings but a compromise had to be made between practicalities and the need for statistical precision.

6.3d. Data collection measures

The York Assessment of Reading for Comprehension (YARC) primary version provided standardised, meaningful, graded, fiction and non-fiction texts with two parallel sets of passages, which enabled comparison of the two fonts and was aimed at the KS2 age group of the study. There was also an online scoring tool, which avoided long manual calculations and resulting errors. By using the SWRT as a placement test each participant was presented with text matching their reading ability so they neither struggled unduly to read the passages nor read them so fluently that there were no errors to analyse. Pijpker (2013) also matched the reading level of participants to the level of the text that they were given. The findings of that study were reflected in this current study regarding reading accuracy and reading errors i.e. both studies found a beneficial impact of a dyslexia-friendly font on reading accuracy. It could be the case that by ensuring that the participants read texts closely corresponding to their reading ability, nuances in reading accuracy could be revealed in the form of significantly different reading accuracy scores. If participants have to read text of a more broadly matched difficulty level then the variation in reading accuracy may not be as quantifiable. None of the other studies outlined in Table 3.6 used texts with a difficulty level corresponding to measured ability for each child, which could lead to results being confounded by text being of an unsuitable level.

The standardisation of the YARC passages ensured that the texts were comparable. This avoided any self-penned texts being incomparable because of variation in difficulty such as reported by Powell and Trice (2019), who found a limitation of findings between their texts due to one text being revealed as easier than the others during analysis.

6.4 Future research.

To ascertain the generalisability of this study it would be necessary to repeat it with the same methodology but a different cohort of participants. In this way, it could be seen whether the results were replicated. Further studies could include participants assigned to the Dyslexia group who have all been tested prior to inclusion using the same testing regime. This would ensure that all had a similar level of reading difficulties and would avoid any variation in the identification of difficulties amongst participating schools.

The methodology of the study could be varied in future to include the addition of a condensed version of OpenDyslexic font . This may serve to give a clearer picture of the role of spacing. It may also be feasible to replicate other extant studies. Presenting the fonts as lists of words (Kuster et al., 2018; Wery & Diliberto, 2016) may indicate how layout of text alongside font design impacts reading performance.

Comparing OpenDyslexic with a different control font e.g. Times New Roman may also influence findings (Wery & Diliberto, 2016). An additional font could show whether test results are affected in the same way as the current study and demonstrate the positive impact of OpenDyslexic or a change of control font might contradict the current findings to demonstrate no significant benefit to OpenDyslexic. A comparison could also be made between OpenDyslexic and one of the other dyslexia-friendly fonts such as Dyslexie or Easy Reading. This would allow investigation as to whether the impact of these fonts compares equally with each other.

6.5 Implications for professional practice

The rationale for this study was to investigate whether a dyslexia-friendly font should be recommended to schools based on research and quantifiable findings. The findings indicate that the use of OpenDyslexic font has a beneficial impact on reading outcomes for pupils in KS2 as demonstrated by better test scores for reading accuracy and reading rate in comparison to a more standard font such as Arial. Thus, passing on this information to schools, parents and pupils is validated.

The most fundamental point regarding this study that needs to be emphasised to schools, parents and pupils concerns what dyslexia is and how font design links to supporting young readers with traits of dyslexia. This is demonstrated by clarifying what dyslexia is not i.e. it is not a visual problem or “word blindness”. There may be some visual confusion linked to comorbid scotopic sensitivity but dyslexia stems primarily from a phonological deficit, the origins of which lie outside the visual system (Evans, 2001; Seidenberg, 2017). None of the definitions of dyslexia indicates a visual component (Table 2.2), so the link between phonological difficulties and the role of font design needs to be outlined. The connection between font design and dyslexia is through legibility. If a particular font makes a text more legible then letter recognition is less cognitively demanding and more cognitive capacity is available to retrieve phonological representations (Thiessen et al., 2015). The effort needed to process text is reduced and the outcome could be more accurate and/or faster reading (Van Someren, 2013).

If a font has distinctive letters with unique features, then young readers do not have to search as hard for distinguishing features (Fiset et al., 2008). They are able to identify graphemes more easily, which they then link to phonemes when decoding. As stated by Willingham (2017) inexperienced readers may search unpredictably for those critical, distinguishing features, so if KS2 pupils have a font with clear letters then identifying crucial elements for letter identification may be less arbitrary. This makes the text more legible overall and improves decoding. As the identification of letters is necessary for all developing readers, the use of a font such as OpenDyslexic with unique letter forms for each letter can be beneficial for all young readers.

As the findings of this study support the idea that a clear font can assist both pupils with dyslexia and typically developing readers, it should be communicated to colleagues, parents and pupils that all pupils can be supported by introducing a dyslexia-friendly font such as OpenDyslexic. This means that if a school introduced OpenDyslexic font universally it would not be detrimental to typically developing readers and would also avoid any stigma of children with dyslexia being given different resources. It would also support any struggling readers who have not been identified as having dyslexia specifically but who would still be supported through a general policy of using OpenDyslexic for resources and displays.

If schools did not wish to implement a whole school approach then OpenDyslexic font could be used to support ongoing interventions. It should be emphasised that adopting OpenDyslexic font is not a substitution for a comprehensive programme of support. OpenDyslexic should be regarded as a compensatory tool for reducing the gap between the reading test scores of pupils with dyslexia and typically developing readers (Bachman and Mengheri, 2018). Parents and pupils especially should be advised that OpenDyslexic font is not a cure for Dyslexia, which was implied in the press when the font was first available. If using OpenDyslexic has the effect of a pupil achieving improved test scores or greater progress with their reading, they still have dyslexia but their difficulties are being ameliorated by text being presented in a different font.

It should be made clear that as every person with dyslexia is different, the impact of OpenDyslexic font varies from child to child. The overall results of this study indicate that the font had a significant effect. However, the impact was considerable on the test scores of some individuals but minimal on the scores of others. Thus, it should not be expected that the font is a panacea. Wery and Diliberto (2016) warn that expectations should be managed to prevent a child, who may have experienced many struggles related to learning to read, from being discouraged further by an additional failure if the font is shown not to impact their reading performance. The use of the font should be prefaced with a proviso that it is another tool to try but failure is not implied if it proves not to be effective for a particular individual and, regardless of the outcome, intervention support will be ongoing. Teachers should also realise that it is not a reflection on their professionalism if the font has little or no effect. They have tried a new approach that in many cases reaps rewards but not for all. All pupils are different and so if the use of OpenDyslexic font is beneficial then it is an impetus that can be built upon (Zikl et al., 2015).

Motivation to read could be affected by a change in font which could impact reading progress. If a pupil feels valued and supported by being able to read in a different font, then maybe they will be more inclined to practice their reading, especially if they then perceive that they are being more successful. Bosman (2014) indicates that font can act as a placebo whereby a child thinks that their reading is improving, which has the effect of building motivation, confidence and then achievement. The best intervention for struggling readers is to read more (Felici, 2012), so using OpenDyslexic font could also be beneficial if only acting on a particular child through a placebo effect.

Although this study concentrated on presenting texts as print on paper, the easiest way to alter font design is on electronic devices. OpenDyslexic is free to download so does not cause an extra expense for cash-strapped schools. Pupils could try out this font on a device in school to gauge its efficacy. If necessary then resources could be prepared for them using OpenDyslexic. Children can be involved in the choice of font so then they have more agency when using technology, as changing font design or size is straightforward.

OpenDyslexic font has been shown to impact test scores but most external testing is via officially produced test materials. These cannot normally be adapted unless exceptional access arrangements are agreed upon. Examination boards etc may not have considered the role of font on reading performance but this could be a discussion for the future.

There is no consensus about the visual attributes of text that are best for beginner readers (Walker et al., 2018). Many schools link a chosen font to their handwriting policy. Often cursive font is used for all resources to provide consistency. Parallel to this should be a consideration of how a dyslexia-friendly font might impact reading results. This study might stimulate a debate as to whether adopting OpenDyslexic font for reading might outweigh the benefit of always using a cursive font, aimed at the improvement of handwriting and presentation. This would depend largely on the schools' development plans. It is not the intention to cause a possible dispute between a SENCo, who wants to improve reading scores and a literacy coordinator, who wants to improve handwriting and spelling. The rationale behind this study is as evidence to underpin advice not to dictate policy.

In addition to schools and parents/carers being involved in children's reading development, there are organisations and charities such as the National Literacy Trust and Coram Beanstalk, who are involved in supporting children's reading. These institutions could also be made aware of the role of font in reading and so could advise their staff and volunteers regarding the choice of texts for young readers and thus consider not just content and reading level but also layout and the impact that font design may be having on reading performance.

6.6 Conclusions

Font design is often a factor that is overlooked when young children learn to read. The role of font design on reading performance may be deemed insignificant. Many teachers may never have considered it. Font could be presumed to have no valuable role other than to record the necessary graphemes and words with no function apart from adding to the overall aesthetic of the page design. This is possibly because well designed print should be inconspicuous (Cullen, 2012). As Warde (1956) described, the crystal goblet that is print allows the content to be enjoyed unhindered, which leads to font design going unnoticed. However, OpenDyslexic could be the key that clears the “fog” that clouds the printed word (Strom, 2003) when a child with traits of dyslexia or a pupil just beginning their journey to fluent reading is presented with a text that they must read and understand. The task of reading could become less arduous for young readers by changing font and improving legibility. As this study shows, an arguably minor variable such as font can have a more proportionate than expected impact on reading performance. Findings illustrate the quantifiable impact of font by showing the occurrence of a significant increase in test scores in reading accuracy and reading rate when text was presented in OpenDyslexic. The expressed opinions of young readers indicate that, generally, they are not resistant to the introduction of a font that has been shown to improve test scores and reading performance. Thus, it is demonstrated that OpenDyslexic has a significant, beneficial impact on the reading performance of Key Stage 2 readers.

References

Adee, S. (2012). Tricksy type: How fonts can mess with your mind. *New Scientist,* 216(2896), 68-69.

Ahr, E., Houdé, O., & Borst, G. (2016). Inhibition of the mirror generalization process in reading in school-aged children. *Journal of Experimental Child Psychology*, 145,157-165.

Alfonsetti, F. (2013). Easy Reading® Font: Highly legible typeface. www.easyreading.it

Alvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in Psychology*, 6:1810

Appert, D. & Truillet, P. (2016). Impact of word presentation for dyslexia. *ASSETS'16,* 265-266.

Applegate, M., Applegate, A. & Modla, V. (2009). She's my best reader; she just can't comprehend: Studying the relationship between fluency and comprehension. *Reading Teacher,* 62(6), 512-521.

Arditi, A., & Cho, J. (2005). Serifs and font legibility. *Vision Research*, 45(23), 2976-2933.

Armbruster, B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read. Kindergarten through grade 3.* ED458 536 Educational Resources Information Center (ERIC)

Bachmann, C., & Mengheri, L., (2018). Dyslexia and fonts: Is a specific font useful? Brain Sciences, 8(5), 89-101.

Barclay, S. (2019). *I wonder what is it like to be Dyslexic?* CPI Print Ltd.

Bates, K. (2004). *Lexie Readable*. www.k-type.com

British Dyslexia Association (BDA). (2009). www.bdadyslexia.org.uk/dyslexic/definitions

Beier, S. (2009). *Typeface legibility: towards defining familiarity*. Unpublished thesis. Royal College of Art.

Beier, S. (2012). *Reading letters: Designing for legibility*. BIS Publishers.

Beier, S., & Dyson, M. (2014). The influence of serifs on ‘h’ and ‘i’: useful knowledge from design-led scientific research. *Visible Language*, 47(3), 74-95.

Beier, S., & Oderkerk, C. (2019). The effect of age and font on reading ability. *Visible Language*, 53(3), 50-68.

Bélanger, N. & Rayner , K. (2015). What eye movements reveal about deaf readers. *Current Directions in Psychological Science*, 24, 220-226.

Benfatto, M., Seimyr, G., Ygge, J., Pansell, T., Rydberg, A. & Jacobson, C.(2016). Screening for Dyslexia using eye tracking during reading *PLoS ONE*, 11(12).e0165508

BERA (2018). *Ethical Guidelines for Educational Research.* British Educational Research Association

Bernard, M., Lida, B., Riley, S., Hackler, T. & Janzen, K. (2002). A comparison of popular online fonts: Which size and type is best? *Usability News,* 4(1).

Bernard, M., Mills, M., Frank, F. & McKown, J. (2001). Which fonts do children prefer to read online? *Usability News,* 3(1).

Bessmans, A. (2016). Typefaces for children’s reading. *Tijdschrift voor mediageschiedenis*, 19(2), 1-9.

Beymer, D., Russell, D., & Orton, P. (2008). An eye tracking study of how font size and type influence online reading. *Conference Proceedings of the 22nd British HCI*

*Group Annual Conference: People and Computers XXII: Culture, Creativity, Interaction - Volume 2.*https:// doi.org/10.1145.1531826.1531831

Bigelow, C. (2019). Typeface features and legibility research. *Vision Research*, 165, 162-172.

Blevins, W. (2017). *A fresh look at phonics: Common causes of failure and seven ingredients for success.* Corwen Literacy.

Boer, C. (2008). *Typeface for readers with dyslexia*. www.dyslexiefont.com

Boets, B., Op de Beeck, H., Vandermosten, M., Scott, S. & Gillebert, C. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science,* 342(6163), 1251-1254.

Bohl, H. & Hoult, S. (2016). *Supporting children with dyslexia (2nd Edition)* Routledge.

Bohm, T. (2019). Letters and symbol misrecognition in highly legible typefaces for general, children, dyslexic, visually impaired and ageing readers (4th edition). www.typography.guru

Bosman, A. (2014). Positieve ervaringen met het lettertype 'Dyslexie": Een placebo effect? [Positive experiences with the font 'Dyslexie': A placebo effect?] *Beter Begeleiden Digitaal,* June.

Bosse, M-L., Tainturier, M-J. & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition.* 104, 198-230.

Bowyer-Crane C. & Snowling, M. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology,* 75(2),189-201.

Brady, S. (2019). The 2003 IDA Definition of Dyslexia: A Call for Changes. *Perspectives on Language and Literacy* 45(1),15–21.

Braun, V. & Clarke, V. (2012). Thematic analysis. In H.Cooper (ed) *APA Handbook of research methods in psychology: Vol 2 Research designs* (pp57- 71) American Psychological Association.

Brem, S., Bach, S., Kucian, K., Kujala, J., Gutthorm, T., Martin, E., Lyytinen, H., Brandeis, D. & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech soung correspondences. *Proceedings of the National Academy of Sciences,* 107(17), 7939-7944.

Breuninger, J. (2019). *Legibility: how to make text convenient to read.* www.uxdesign.cc

Breznitz, Z. (2005). *Fluency in reading*. Taylor & Francis Group.

Broadbent, L. (2018). *OpenDyslexic font: Impact on the reading accuracy and comprehension of key stage 2 readers with dyslexia*. Unpublished IFS Study, University College London, Institute of Education.

Briem, G, (2002). How to arrange text on web pages. In R. Sassoon (ed.) *Computers and Typography 2.* Intellect Ltd.

Brownlee, J. (2014). *What's the difference between a font and a typeface? Three minute read.* www.fastcompany.com

Budiu, R and Moran, K (2021) How many participants for quantitative usability studies: a summary of sample-size recommendations www.nngroup.com

Burgess, T. (2012). *Open dyslexic, does the new font help?* www.examiner.com/article/opendyslexia-font-help-review=by=teacher=and-students

Caccamise, D., Snyder, L. & Kintsch. E. (2008) Constructivist theory and the situation model: Relevance to future assessment of reading comprehension. In C. Collins Block & S. Parris (eds.) *Comprehension Instruction: Research-based best practices,* 2nd ed. pp.80-87. The Guilford Press

Çağlayan, A. (2009). *Designing a new typeface for children*. Unpublished thesis, Izmir University of Economics, Turkey.

Cain, K. (2010*). Reading Development and Difficulties*. BPS Blackwell/John Wiley.

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

Carroll,J., Solity, J. & Shapiro, L. (2015). Predicting dyslexia using prereading skills: the role of sensorimotor and cognitive abilities. *Journal of Child Psychology* 57(6), 750-758.

Casini, L., Pech-Georgel, C. & Ziegler, J. (2017). It's about time: revisiting temporal processing deficits in dyslexia. *Developmental science, - Wiley Online Library*

Castles, A. & Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind and Language,* 29(3), 270-285.

Changizi, M., Qiong,Z., Ye, H. & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist* 167(5), 117-139.

Clarke, T. (2018). *Best UX practices for line spacing. UX design- research and insights*. www.justinmind.com

Clay, M. (1991). *Becoming literate: The construction of inner control.* Heinemann Education.

Clay, M. (1993). *Reading recovery: A guidebook for teachers in training.* Heinemann Education.

Clay, M. (2000). *Running records for classroom teachers.* Heinemann Education.

Clay, M. (2001). *Change over time in children's literacy development.* Heinemann Education

Coch, D. (2010). Constructing a reading brain. In D. Sousa (Ed.), *Mind, brain and education* (pp.138-161). Solution Tree Press.

Colenbrander, D., Nickels, L. & Kohnen, S. (2016). Similar but different: differences in comprehension diagnosis on the Neale Analysis of Reading Ability and the York Assessment of Reading for Comprehension. *Journal of Research in Reading, 40(4), 403-419.*

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud *Psychological Review* 108 (1), 204-256.

Cooke, A. (2001). Critical response to dyslexia, literacy and psychological assessment: a view from the chalk face. *Dyslexia* 7(1), 47-52.

Cornelissen, P., Richardson, A., Mason, A., Fowler, S. & Stein, J. (1995). Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics and controls.Vision Research, 35(10), 1483-1494.

Cromley, J., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension.
Journal of Educational Psychology, 99(2), 311-325.

Cullen, K. (2012). *Design elements: Typography fundamentals*.
Rockport Publishers.

Davenport, M. (2002). *Miscues not mistakes: Reading assessment in the classroom*. Heinemann.

Davidow, A. (2002). Computer screens are not like paper: typography on the web. In R. Sassoon (ed.) *Computers and Typography 2.* Intellect Ltd.

Dehaene, S. (2009). *Reading in the brain*. Viking.

Dehaene, S. (2014). Reading in the brain revised and extended: Response to comments. *Mind & Language*, 29(3), 320-335.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384-398.

Dehaene,S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15(6), 254-262.

Dehaene,S., Cohen, L., Morais, J.& Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquistion. *Nature reviews Neuroscience* 16 234-244

De Leeuw, R. (2010). *Special font for dyslexia?* Masters Thesis, University of Twente, NL. www.essay.uTwente.nl

Del Real Garcia, M. (2014). Dixy: A connection between the typographic appearance and the brain of children. *Journal of Literature and Art Studies, 4*(1), 55-61.

De Luca, M., Borelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language,* 80, 617-626.

De Soto, D., (2014). *Know your onions: Graphic design*. BIS Publishers.

Diemand-Yauman, C, Oppenheimer, D. & Vaughan, E. (2011). Fortune favors the bold and the italicised; Effects of disfluency on educational outcomes. *Cognition* 118(1), 111-115.

Dockett, S., Perry, B. & Kearney, E. (2012). Promoting children's informed assent in research participation . *International Journal of Qualitative Studies in Education* 26 (7), 802-828.

Draper, S. (2016). Effect size www.psy.gla.ac.uk/~steve/effects

Dressler, E. (2019). Understanding the effects of font type on reading comprehension/memory under time-constraints. *Theses/Capstones/Creative Projects.70.*

Duranovic, M., Senka, S. & Babic-Gavric, B. (2018). Influence of increased letter sopacing and font ttype on the reading ability of dyslexic children. *Annals of Dyslexia*, 68, 218-228.

Dyslexia Foundation of NewZealand (2008) *Recognising Dyslexia* www. Dyslexiafoundation.org.nz

Dyson, M., & Beier, S. (2016). Investigating typographic differentiation: italics are more subtle than bold for emphasis. *Information Design Journal*, 22(1), 3-18.

Dyson,M., & Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. *International journal of Human-Computer studies*, 54, 585-612.

Eden, G. (2017). *Is there a certain font that works best for kids with dyslexia?* www.understood.org/en/learning

Ehri, L.(2005). Development of sight word reading: Phases and findings. In M.Snowling & C. Hulme (Eds.) *The science of reading: a handbook* (pp.135-154). Blackwell Publishing.

Elliot, J. (2020) It's time to be scientific about dyslexia. *Reading Research Quarterly* 55 (51), 561-575.

Elliot, J., & Grigorenko, E. (2014). *The dyslexia debate.* Cambridge University Press.

Epstein, V. (2016) Dyslexia-Does font really matter? www.learningally.org/blog

Evans, B. (2001). *Dyslexia and Vision*. Whurr Publishers Ltd.

Farley, J. (2010). Typography: readability and Legibility (part 1). www.sitepoint.com/blogs

Felici, J. (2012). *Towards a typographic prescription for dyslexia* www.creativepro.com

Fiset, D., Blais, C., Éthier-Majcher., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, 19 (11), 1161-1168.

Foster, H. (2007). *The Single Word Reading Test 6-16*. London: NFER Nelson Publishing Company Ltd.

Franzen, L, Stark, Z, Johnson, A. (2019). The dyslexia font OpenDyslexic facilitates visual processing of text and improves reading comprehension in adult dyslexia. *Annals of Eye Science,* 4, AB004

Franzen, L, Stark, Z, Johnson, A. (2021) Individuals with dyslexia use a different visual sampling strategy to read text . *Scientific reports: nature portfolio 11. 6449.*

French, M., Blood, A., Bright, N., Futak, D., Grohmann, M., Hasthorpe, A., Heritage, J., Poland, R., Reece, S. and Tabor, J. (2013). Changing fonts in education: How the benefits vary with ability and dyslexia. *The Journal of Educational Research,* 106(4), 301-304.

Frensch, N. (2003). *Read Regular*. www.readregular.com

Gabrieli, J., Christodoulou, J., Loughlin, T., & Eddy, M. (2010). The reading brain. In D.Sousa (Ed.), *Mind, brain and education* (pp.113-136).
Solution Tree Press.

Galliussi, J., Perondi, L., Chia, G., Gerbino, W., & Pernardis, P. (2020).
Inter-letter spacing, inter-word spacing, and font with dyslexia- friendly features: testing text readability in people with and without dyslexia. *Annals of Dyslexia* 70, 141-152.

Garfield, S. (2011). *Just my type: a book about fonts*. Profile Books.

Gathercole, S., & Baddeley, A. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, 8, 259-272.

Gendelman, V. (2015). *Font psychology: How typefaces hack our brains*. www.companyfolders.com/blog/font-psychology-how-typefaces-hack-our-brains

Gentry, J., & Ouellette, G. (2019). *Brain words: How the science of reading informs teaching.* Stenhouse Publishers.

Georgiou, G., Papadopulos, T., Zarouna, E. & Parrila, R. (2012) Are auditory and visual processing deficits related to developmental dyslexia? *Dyslexia* 18 (2).110-129.

Giraud, A-L. & Ramus, F. (2013). Neurogenetics and auditory processing in developmental dyslexia. *Current Opinion in Neurobiology,* 23(1), 37-42.

Goodale, M., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.

Goodman, K. (1969). Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly,* 5, 9-33.

Gonzalez, A. (2012).*Does this font truly help dyslexic readers to read?* www.opendyslexicfont.com

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

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review,101*(3), 371–395.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence fron the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102, 167-181.

Hakvoort, B., van den Boer, M., Leenaars, T., Bos ,P., & Tijms, J. (2017). Improvements in reading accuracy as a result of increased interletter spacing are not specific to children with dyslexia. *Journal of Experimental Child Psychology*, 164, 101-116.

Haley, A. (2017). *It's about legibility*. www.fonts.com

Harmey, S. & Kabuto, B. (2018). Metatheoretical differences between running records and miscue analysis: Implications for analysis of oral reading behaviours. *Research in the Teaching of English,* 53(1), 11-33.

Heller, S. & Anderson, G. (2017). *Type tells tales*. Thames & Hudson Ltd.

Herrmann, R. (2011). *What makes letters legible* www.opentype.info/blog/typography

Hillier, R. (2006). A typeface for the adult dyslexic reader. www.robsfont.com

Hodgkinson, T., & Small, D. (2018). Orienting the map: Where K to12 teachers stand in relation to text complexity. *Literacy research and instruction,* 57(4), 369-386.

Hughes, L., & Wilkins, A. (2000). Typography in children's reading schemes may be suboptimal: Evidence from measures of reading rate. *Journal of Research in Reading, 23(3),* 314-324.

Hughes, L., & Wilkins, A. (2002). Reading at a distance: Implications for the design of text in children's big books. *British Journal of Educational Psychology*, 72, 213-226.

Hulme, C. (1988). The implausibility of low-level visual deficits as a cause of children's reading difficulties. *Cognitive Neuropsychology,* 5, 369-374.

Hunziker, H. (1998). Typographic work of Adrian Frutiger. *Serif: the magazine of type & typography*, 6, 32-43.

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

Joo, S., White, A., Strodtman, D., & Yeatman, J. (2018). Optimizing text for an individuals visual system: The contribution of visual crowding to reading difficulties. *Cortex*,103, 291-301.

Joseph, H., Liversedge, S.,Blythe, H., White. S., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49, 2078-2086.

Joseph,H. & Powell,D. (2022). Does a specialist typeface affect how fluently children with and without dyslexia process letters, words, and passages? *Dyslexia,* 28(4), 1-23.

Josephson, S. (2008). Keeping your readers' eyes on the screen: An eye-tracking study comparing sans serif and serif typefaces. *Visual Communication Quarterly*, 15(1-2), 67-79.

Joshi, R., & Aaron, P. (2008). Assessment of literacy performance based on the componential model of reading.
In G. Reid, A. Fawcett, F. Manis, L.Siegel. (Eds.), *The sage handbook of dyslexia* (pp.268-290). Sage Publications Ltd.

Katzir,T., Hershko, S., & Halamish, V. (2013). The effect of font size on reading comprehension on second and fifth grade children: Bigger is not always better. *PLOS One* 8(9),1-8.

Kelly,K & Phillips, S, (2016). *Teaching literacy to learners with dyslexia 2nd edition* Sage Publications Ltd

Kendeou, P., Bohn-Gettler, C., White, M., & van den Broek, P. (2008). Children's inference generation across different media.
Journal of Reading 31(3), 259-272.

Khataee, E., & Davoudi, M. (2018). The role of cultural schemata in inferential reading comprehension: An investigation in the Iranian EFL context.
*ASEAN Journal of Teaching & Learning in Higher Education,*10(2),11-27.

King, S. (2018) "Sir, it works!" The Dyslexie font. *The School Librarian* 66(2), 72-74.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2),163–182.

Kintsch, W. & Rawson, K. (2005). Comprehension. In M. Snowling & C. Hulme (eds.) *The science of reading: a handbook* (pp209-226). Blackwell Publishing

Kirby, J., & Savage, R. (2008). Can the simple view deal with the complexities of reading? *Literacy*, 42(2), 75-82.

Kispal, A. (2008). *Effective teaching of inference skills for reading*. National Foundation for Educational Research, Department for Children, Schools and Families.

Kuhn, M., Schwanenflugel, P., & Meisinger, E. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody and definitions of fluency. *Reading Research Quarterly*, 45(2), 230-251.

Kuster, S., Weerdenburg, M., Gompel, M., & Bosman, A. (2018). Dyslexie font does not benefit reading in children with or without dyslexia. *Annals of Dyslexia*, 68(1), 25-42.

Lahman, M., Rodriguez, K., Moses, L., Griffin, K., Mendoza, B., & Yacoub, W. (2015). A rose by any other name is still a rose? Problematizing pseudonyms in research. *Qualitative Inquiry*, 21(5), 445-453.

Lervåg, A. & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science,* 20(8), 1040-1048.

Lilienfeld, S., Steven, J., & Namy, L. (2018). *Psychology: From inquiry to understanding.* Pearson.

Livingstone, M., Rosen, G., Drislane, F. & Calaburda, A. (1991). Physiological evidence for a magnocellular deficit in developmental dyslexia. *Proceedings of the New York Academy of Science*, 88, 7943-7947.

Lobier, M. & Valdois, S. (2015). Visual attention deficits in developmental dyslexia cannot be ascribed solely to poor reading experience. *Nature Reviews Neuroscience,*16, 225-226.

Loyd, J. (2013). *Typographic readability and legibility*. www.webdesign.tutsplus.com

Luna, P. (2018). *Typography: A very short introduction*. Oxford University Press.

Lupton, E. (2014). *Thinking with type: A critical guide for designers, writers, editors & students (expanded edition).* Princeton Architectural Press.

Lyon, G., Shaywitz, S., & Shaywitz, B. (2003). A definition of dyslexia. *Annals of Dyslexia* 53, 1-14.

Magombe, A. (2011). *Research into typographic design for young people*. www.issuu.com

Martinet, C., Valdois, S. & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, 91(2),11-22.

Marinus, E., Mostard, M., Segers, E., Schubert, T., Madelaine, A., & Wheldall, K. (2016). A special font for people with dyslexia: Does it work and, if so, why? *Dyslexia*, 22, 233-244.

Martelli, M., Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision*, 9(4), 1-18.

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder,P., Steinhausen, H., & Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *NeuroImage* 33(2), 749-758.

Melby-Lervåg, M., Halaas Lyster, S-A. & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta- analytic review. *Psychological Bulletin,* 138(2), 322-352.

Mileta, N. (2014). *Setting body text for comfortable reading*. (March, 5) www.magazinedesigning.com

Moll, K., & Jones, M. (2013). Naming Fluency in Dyslexic and Nondyslexic Readers: Differential Effects of Visual Crowding in Foveal, Parafoveal, and Peripheral Vision. *Quarterly Journal of Experimental Psychology, 66*, 2085 – 2091.

Monzalvo, K., Fluss, J., Billard, C., Dehaene, S.& Dehaene-Lambertz, G. (2012). Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status, *NeuroImage,* 61 (1), 258-274.

Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes childrens spoken language network. *Brain & Language*, 127(3), 356-365.

Moret-Tatay, C., & Perea, M. (2011). Do serifs provide an advantage in the recognition of written words? *Journal of Cognitive Psychology*, 23(5), 619-624.

Morley, M. (2018). Can fonts really help those with dyslexia? www.eyeondesign.aiga.org/can-fonts-really-help-those-with-dyslexia

Morris , R., Aquilante, K., Bigelow, C., & Yager, D. (2002). Serifs slow reading at very small sizes, but don't matter at larger sizes. *Paper presented at the SID Digest of Technical papers symposium, Boston, USA.*

Morrow, V., (2008). Ethical dilemmas in research with children and young people about their social environments, *Children's Geographies,* 6(1), 49-61.

Muter, V. (2021) *Understanding and supporting children with literacy difficulties: An evidence-based guide for practitioners.* Jessica Kingsley Publishers

Neale, M. (1989) *The Neale Analysis of Reading Ability*, (Revised British edition.) NFER-Nelson.

Neale, M. (1999). *The Neale Analysis of Reading Ability*. (Third edition). ACER.

Neisser, U. (1967). *Cognitive Psychology (Classic Edition).*Psychology Press and Routledge Classic Editions.

Nicolson, R. & Fawcett, A. (2008). *Dyslexia, learning and the brain.* The MIT Press.

Nicolson, R., Fawcett, A. & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in Neurosciences,* 24(9), 508-511.

Oakhill, J., & Cain, K. (2018). Children's problems with inference making: causes and consequences*. Bulletin of Educational Psychology,* 49(4), 683-699.

Oakhill, J., Berenhaus, M., & Cain, K., (2015). Children's reading comprehension and comprehension difficulties. In A. Pollatsek & R. Treiman. (Eds.). *Oxford Handbook of Reading*. Oxford Handbooks Online

O'Brien, B., Mansfield, S., & Legge, G. (2005). The effect of print size on reading speed in dyslexia. *Journal of Research in Reading*, 28(3). 332-349.

Onwuegbuzie, A. and Leech, N. (2005). On becoming a pragmatic researcher: Importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology*, 8(5), 375-387.

Opsteegh, M. (2010). Writing in type: The art of choosing typefaces that work. *Intercom*, 57(3),12-16.

Ossen, A. (2012). Effecten van het lettertype Dyslexie in het basisonderwijs- een praktijkonderzoek. [Effects of the font Dyslexie in primary education -a practical investigation. *Beter Begeleiden Digitaal* June

Parker, A., Kirkby, J., & Slattery, T. (2020). Undersweep fixations during reading in adults and children. *Journal of Experimental Child Psychology*, 192, 1-23.

Pastel, B. (2011). Are children sensitive to typeface? *Dissertation, University of Reading*

Pattison, D. (2020) What fonts for children's books? www.indiekidsbooks.com

Pelli, D., Waugh, S., Martelli, M., Crutch, S., Primativo, S., Yong, K., Rhodes, M., Yee, K., Wu, X., Famira, H., & Yiltiz, H. (2016). A clinical test for visual crowding . F1000 research.htpps:// doi.org/ 10.12688/f100research.7835.1

Pennington, B.(2006). From single to multiple deficit models of developmental disorders. *Cognition* 101, 385-413.

Pennington, B., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D., Samuelsson, S., Byrne, B., Willcutt, E. & Olson, R. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology, 121*(1), 212–224.

Perea, M. & Gomez, P. (2012). Subtle increases in interletter spacing facilitate the encoding of words during normal reading. PLOS One, 7(10).

Perea, M., Panadero, V., Moret-Tatay, C. & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction*, 22, 420-430.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading,* 11(4), 357-383.

Phillips, S., & Kelly, K. (2018) *Assessment of learners with dyslexic type difficulties* (2nd Ed.). Sage.

Phinney, T., & Colabucci, L. (2010). The best font for the job. *Children & Libraries*, 8(3),17-26.

Pijpker, C. (2013). *Reading performance of dyslexics with a special font and a coloured background*. University of Twente, NL. www.essay.utwente.nl

Plaut, D. (2005). Connectionist approaches to reading. In M.Snowling & C. Hulme (Eds.) *The science of reading: a handbook* (pp.24-38). Blackwell Publishing.

Powell. S., & Trice, A. (2019). The impact of a specialised font on the reading performance of elementary children with reading disability. *Contemporary School Psychology*, 24, 34-40.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Lawrence Erlbaum Associates, Inc.

Quigley, A. (2020). *Closing the reading gap*. Routledge.

Raban, B. (1982). Text display effects on the fluency of young readers. *Journal of Reading Research,* 5(1), 7-28.

Ramus, F. (2014) Should there really be a Dyslexia debate? *Brain* 137(12), 3371-3374.

Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S. & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain,* 126, 841-865.

Rasinski, T. & Cheesman Smith, M. (2018). *The megabook of fluency: Strategies and texts to engage all readers*. Scholastic.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin,* 124(3), 372-422.

Rayner, K., Slattery, T., & Bélanger, N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, 17(6), 834-839.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*(1),16–20.

Reid, G. (2016). *Dyslexia: A practitioner's handbook. (5th ed.)* John Wiley & Sons Ltd.

Rello, L and Baeza-Yates, R. (2013). Good fonts for dyslexia. *ASSSETS, 2013,* www.luzrello.com/Publications_files/assets2013.pdf

Rello, L & Baeza-Yates, R. (2017). How to present more readable text for people with dyslexia. *Universal Access in the Information Society*, 16, 29-49.

Reynolds, L. & Walker, S. (2003). Serifs, sans serifs, and infant characters in children's reading books. *Information Design Journal*, 11(3),106-122.

Reynolds, L. & Walker, S. (2004). You can't see what the words say: word spacing and letter spacing in children's reading books.
Journal of Research in Reading, 1(27), 87-98.

Reynolds, L., Walker, S. & Duncan, A. (2006). Children's responses to line spacing in early reading books or 'holes to tell which line you're on'.
Visible Language, 40(3), 246-267.

Ripoll, J., (2015). Font legibility in first year primary students. *Journal for the study of education and development,* 38(3), 600-616.

Robson, C. (2011). *Real World Research : a Resource for Users of Social Research Methods in Applied Settings*. Wiley.

Rodgers, E., D'Agostino, J., Kelly, R., & Mikita, C. (2018) Oral reading accuracy: findings and implications from recent research.
The Reading Teacher, 72(2),149-157.

Rose, J. (2009). Identifying and teaching children and young people with dyslexia and literacy difficulties. *DfE.* www.webarchive.nationalarchives.gov.uk

Rummer,R., Schweppe, J. & Schwede, A. (2016) Fortune is fickle: null-effects of disfluency on learning outcomes. *Metacognition Learning* 11, 57-70.

Sadoski, M. (2008). Dual coding theory: reading comprehension and beyond.
In C. Collins Block & S. Parris (eds.) *Comprehension Instruction: Research-based best practices,* 2nd ed. pp.38-49. The Guilford Press

Saha, N. (2020) Do dyslexia friendly fonts actually work? www.readingreserchrecap.substack.com

Sassoon, R. (1993). *Computers and Typography.* Oxford: Intellect Books.

Savage, R., Stuart, M. & Hill, V. (2001). The role of scaffolding errors in reading development: Evidence from longitudinal and a correlational study. *British Journal of Educational Psychology*, 71(1),1-13.

Schneps, M., Thomson, J., Sonnert, G., Pomplum, M., Chen, C. & Heffner-Wong, A. (2013a). Shorter lines facilitate reading in those who struggle. PLOS ONE, 8(8).

Schneps, M., Thomson, J., Chen, C., Sonnert, G & Pomplun, M. (2013b). E-readers are more effective than paper for some with dyslexia. PLOSONE 8(9).

Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it.* Hatchette Book Group.

Seymour, P. (2005). Early Reading Development in European Orthographies. In M. Snowling & C. Hulme (Eds.) *The science of reading: A handbook* (pp. 296–315). Blackwell Publishing.

Shafig, A. (2015). *Introduction to typography*. www.academia.edu

Shaikh, A. (2005). The effects of line length on reading online news. *Usability news* 7(2), 1-4.

Shanahan, T., Fisher, D. & Frey, N. (2012). The challenge of challenging text. *Educational Leadership* 69(6), 58-62.

Shovman, M. & Ahissar, M. (2006). Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research,* 46, 3514-3525.

Siegel, L. & Hurford, D. (2019) The case against discrepancy models in the evaluation of dyslexia. *Perspectives on Language and Literacy* 45(1) 23-28.

Skottun, B. (2005). Magnocellular reading and dyslexia. *Vision Research,* 45, 113-134.

Skottun, B., & Skoyles, J. (2012). Interletter spacing and dyslexia.
PNAS, 44(109), https:// doi.org/ 10.1073/pnas.1212877109

Slattery, T., Yates, M., & Angele, B. (2016). Interword and interletter spacing effects during reading revisited: Interactions with word and font characteristics. *Journal of Experimental Psychology:Applied*. 22(4), 406-422.

Smith, B. (2016). *Does font really matter to children with dyslexia?* In Learning and Language Difficulties. www.aulexic.com.au

Smythe, I. (2010). *Dyslexia in the digital age. Making IT work*. Continuum International Publishing Group.

Snowling, M. (2000). *Dyslexia* (*2nd Edition*). Blackwell Publishing.

Snowling M. (2013). Early identification and interventions for dyslexia: A contemporary view. *Journal of Research in Special Educational Needs,* 13(1), 7-14.

Snowling, M., Cain, K., Nation, K. & Oakhill, J. (2009a). *Reading comprehension: nature, assessment and teaching*. ESRC seminar series, Reading comprehension from theory to practice, eprints.lancs.ac.uk

Snowling, M., Stothard, S., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K., & Hulme, C. (2009b). *York Assessment of Reading for Comprehension,* GL Assessment.

Sousa, D. (2014). *How the brain learns to read* (2nd ed.). Corwin.

Sperling, A., Lu, Z., Manis, F. & Seidenberg, M. (2006). Motion-perception deficits and reading impairment: It's the noise, not the motion. *Psychological Science*,17 (12), 1047-1053.

Spinelli, D., De Luca, M., Judica., A. & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex*, 38, 179-200.

Stanovich, K. (1986).Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly,* 21(4), 360-407.

Stein, J. (2014). Dyslexia: The role of vision and visual attention. *Current developmental disorders reports,* 1(4), 267-280.

Stein, J & Walsh, V. (1997). To see but not to read: The magnocellular theory of dyslexia. *Trends in Neuroscience,* 20,147-152.

Stinson, M., & Elnar, R. (2016). *Choosing and pairing typefaces for cookbooks.* www.blog.typekit.com/TypeEd

Stoodley, C. & Stein, J. (2013). Cerebellar function in developmental dyslexia. *The Cerebellum,* 12(2), 267-276.

Strizver, I. (2011). *Typography for children*. www.fonts.com

Strizver, I. (2014). *Six typography terms that get confused a lot*. www.creative.live.com/blog/category/topics/art-design

Strizver, I. (2018). *Legibility an readability: What's the difference?* www.creativepro/design

Strom, K. (2003) Reading Scott Makela: the subversion of dyslexic deconstruction. *Design Issues*, 19(2), 5-16.

Stuart, M., & Coltheart, M. (1988). Does reading develop in a series of stages? *Cognition*, 30(2),139-181.

Tashakkori, A. & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioural Research*, Sage.

Tannock, R. (2014). DSM-5 changes in diagnostic criteria for specific learning disabilities (SLD): What are the implications? *International Dyslexia Association* www.dyslexiahelp.umich.edu

Taylor, J. & Perfetti, C. (2016). Eye movements reveal readers' lexical quality and reading experience. *Reading and Writing: An Interdisciplinary Journal*, 29(6), 1069-1103.

Thiessen, M. & Dyson,M. (2010). Clearer and Better :Preference of children with reading difficulties for the typography and illustration in literacy materials. *The International Journal of Learning* 16(12), 365- 384.

Thiessen, M., Kohler, M., Churches, O., Coussens, S. & Keage, H. (2015). Brainy type: a look at how the brain processes typographic information. *Visible Language,* 49(1/2), 175-189.

Thomson, J., Barzillai, M., van den Broek, P., & Schroeder, S. (2018). Learning to read in a digital world: Discussion. In M. Barzillai, J. Thomson, S. Schroeder, P. van den Broek, (eds.). *Learning to Read in a Digital World* (pp236-254). John Benjamins Publishing Company.

Tiffin-Richards, S., & Schroeders, S. (2018). The development of wrap-up processes in text reading: A study of children's eye movements. *Journal of Experimental Psychology, Memory and Cognition*, 44(7), 1051-1063.

Tracy, W. (1988*). The typographic scene*. Gordon Fraser.

University of Lille, The Dyslexie font, Eye tracking research, (English translation) www. DysleixeFont.com

van Bergen, E., van der Leij, A., & de Jong, P. (2014) The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience*, 8 (346), 1-11.

van den Boer, M., & Hakvoort, B. (2015). Default spacing is the optimal spacing for word reading. *Quarterly Journal of Experimental Psychology*, 4(68), 697-709.

van Someren, L. (2013). Welke aanwijzingen in de verklaringstheorieën van dyslexie zijn te vinden voor de bevinding dat dyslectici meer accuraat lezen als woorden worden aangeboden in het lettertype “Dyslexie”?
[What evidence can be identified in the theories of dyslexia for the finding that those with dyslexia can read more accurately when words are presented in the font “Dyslexie”?] *Unpublished thesis, Universiteit van Amsterdam.*

Vellutino, F., Fletcher, J., Snowling, M. & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology & Psychiatry*, 45(1), 2-40.

Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. In G. Underwood (ed.) *Cognitive processes in eye guidance* (pp. 1-32). Oxford University Press.

Walker, S. (2001). *Typography & language in everyday life: prescriptions and practices*. Routledge.

Walker, S., (2005) *The songs the letters sing: typography and children’s reading*. National Centre for Language and Literacy.

Walker, S., Black, A., Bessemans, A., Bormans, K., Renckens, M., & Barratt, M. (2018). Designing digital texts for beginner readers: performance, practice and process. In M. Barzillai, J. Thomson, S. Schroeder, P. van den Broek, (eds.). *Learning to Read in a Digital World* (pp31-56). John Benjamins Publishing Company.

Walker, S. & Reynolds, L. (2003). Serifs, sans serifs and infant characters in children’s reading books. *Information Design Journal,* 11(2/3), 106-122.

Warde, B. (1956). *The crystal goblet: sixteen essays on typography*. World Publishing Company.

Wery, J. & Diliberto, J. (2016). The effect of a specialized dyslexia font, OpenDyslexic, on reading rate and accuracy. *Annals of Dyslexia,* 67(2), 114-127.

Wilkins, A., Cleave, R., Grayson, N. & Wilson, L. (2009). Typography for children may be inappropriately designed. *Journal of Research in Reading*, 32(4), 402-412.

Wilkinson, A. (2005). Man of letters: Onward and upward with the arts. *The New Yorker*, 81(39). 56-65.

Willingham, D. (2017). *The reading mind: A cognitive approach to understanding how the mind reads.* Jossey-Bass.

Wolf, M. & Katzir-Cohen, T. (2001). Reading fluency and its intervention *Scientific Studies of Reading,* 5(3), 211-239.

Wong, I. (2020). An exploratory study to investigate eye movement performance and visual perceptual skills in children with dyslexia.
Asia Pacific Journal of Developmental Differences, 7(1), 27-60.

Woods, R., Davis, K. & Scharff, L. (2005). Effects of typeface and font size on legibility for children. *American Journal of Psychological Research*, 1(1), 86-102.

YARC (2009) York Assessment of Reading For Comprehension (YARC) www.gl-education.com

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

Zender, M. (2019). Letterform legibility and visual perception: A speculation. *Visible Language*, 53(3), 70-91.

Zhang, Y. (2006). The effects of font design characteristics on font legibility. *Unpublished thesis*, Concordia University, Montreal, Canada.

Zhou,S., Jeong, H., & Green, P. (2017). How consistent area the best known readability equations in estimating the readability of design standards?
IEEE Transactions on Professional Communication, 60(1), 97-111.

Zikl, P., Bartošová, I., Višková, K., Havličková, K., Kučirková, A., Navrátilová, J., & Zetková, B. (2015). The possibilities of ICT use for compensation of difficulties with reading in pupils with dyslexia. *Procedia-Social and Behavioural Sciences,* 176, 915-922.

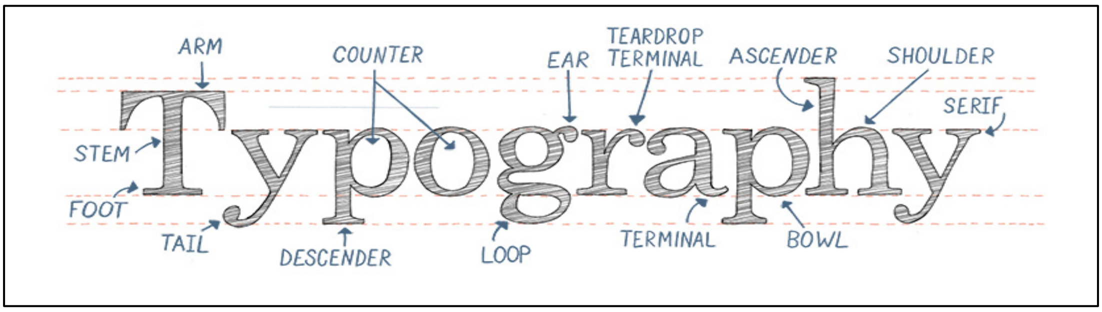
Zorci, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., Bravar, L., George, F., Pech-Georgel, C. & Ziegler, J. (2012a).
Extra-large letter spacing improves reading in dyslexia. *Proceedings of the national academy of science.* 109(28), 11455-11459.

Zorzi, M., Barbiero,C., Facoetti, A., & Ziegler, J. (2012b) Reply to Skottun and Skoyles: Statistical and practical significance of extra-wide letter spacing for dyslexic children. *PNAS*,*Biological sciences* 109(44). 2958-2959.

Appendices

Appendix 1: Glossary of terms & anatomy of a letterform

Diagram to show definitions for individual parts of letterforms (Maulini, 2017)



Arm: A horizontal or upward sloping stroke that does not connect at one end.

Ascender: The portion of lowercase letters that extends above the x-height.

Bowl: The enclosed round part of a letter.

Counter: The counter is the either partially or fully enclosed space of a letter.

Descender: The portion of letters that extends below the baseline.

Ear: A small element that extends from the upper corner of the g or r.

Foot: Part of the stem that rests on the baseline.

Loop: The enclosed counter below the baseline such as the double storey g.

Serif: The small detail attached to the end of a stroke.

Shoulder: A curved part of the letter that extends from the stem.

Spine: The main curved stroke of the letter ‘s’.

Stem: The main vertical stroke of a letter.

Stroke: line element of letter, horizontal, vertical or diagonal; straight or curved.

Tail: The descending stroke on the letter Q or R and g, j, p, q and y.

Terminal: a curved ending to a letter.

Appendix 2:Parent/guardian consent form

Dear Parent/Guardian

My name is Liz Broadbent and I am a qualified teacher studying for a Doctorate in Education at University College London, Institute of Education.

I am researching whether the design of fonts, used by schools and in books, affects how children learn to read. I am especially interested in whether this has an impact on children with and without reading difficulties. I will be studying pupils who are fluent readers, as well as those who find reading a challenge.

This research is important because if it is found that the font design does make a difference, then schools and publishers can take the simple measure of using a certain font to make learning to read an easier task for many children. Up to now there has been little formal research in this area.

I would like to invite your child to take part. If you agree to your child participating, then they will be asked to read short passages tailored to their reading ability, two printed in a standard font and two in a specialist font. They will then answer a few verbal questions about what they have read to measure their comprehension of each passage. They will also be asked their opinion of the two fonts.

If at any time during the reading session your child decides that they do not wish to continue, they can withdraw without needing to give a reason. The sessions will be audio-recorded for analysis but will not be revealed to the school or anyone outside of the project and will be deleted at the end of the study. Data will be kept under the terms of the General Data Protection Regulation (GDPR). Your child and school's identity will not be disclosed in the report I write afterwards, which will include general findings, not information about individuals.

At the end of the study the parents of the participants will be informed of the overall findings.

Further information about the study is available on the accompanying leaflet.

I very much hope that you will consent to your child taking part in this study. Could you please discuss this with your child and provide your written consent by signing and returning the slip to …………………….by……………………….(date).

Yours faithfully,

Mrs Liz Broadbent, MA
University College London, Institute of Education,

✂---

Please fill in, detach and return this section to ……………by……………...

Name of child……………………………………………Class…………………

I agree to the above named taking part in the study.

Signed……………………………………………………..   date…………………

Appendix 3: Parent/ guardian information form

The Research Study

Standard or specialist font?

the quick brown fox jumped over the lazy dog THE QUICK BROWN FOX JUMPED OVER THE LAZY DOG	the quick brown fox jumped over the lazy dog THE QUICK BROWN FOX JUMPED OVER THE LAZY DOG

The aim of the research is to find out if children read and comprehend differently when reading from a standard font (Arial) or a specialist font designed as being dyslexia-friendly (OpenDyslexic). Font designers make claims that are not always backed by research, so I will consider whether font design makes a difference to the reading achievement of young readers or if it has no effect on reading success at all. I will also consider whether the spacing of letters and words makes a difference.

Participants will read passages from a standardised reading comprehension test including passages written in OpenDyslexic font. They will then answer a few questions, verbally, about each passage. I will then be able to see if they have read more easily or more accurately using the standard font or the specialist font and whether the font has had an effect on their understanding of what they have read. The participants will also be asked their opinions of the different fonts and any effect that they feel different fonts have on their reading experience.

I need readers of all abilities and ages to take part, not just those who find reading challenging, so it does not imply that a pupil who is invited to participate has any difficulties. Individual pupils will not be identified and the school will not be named in the research. No one will be able to identify participants when the results are shared. It is hoped that ultimately this study will help the process of learning to read go more smoothly for every young person.

At the end of the study it is hoped that the overall findings can be shared with school staff, pupils who took part and their parents/guardians. It will not be possible to report on results of individuals. The results will show participants how by giving their time and assistance to this research they might support other pupils with learning to read in the future. It might also give pupils themselves greater insight into font choices to facilitate reading.

All data will be protected in line with data protection legislation GDPR and DPA 2018. The controller for this project is University College London (UCL). Further information on how the university uses participant information can be found by accessing www.ucl.ac.uk/legal-services/privacy/ucl-general-research-participant-privacy-notice

Please note that all assurances on confidentiality will be strictly adhered to. Audio recordings of the activities made during this research will be used for analysis only. No one outside the project will be allowed access to the original recordings. They will be stored securely and anonymously. They will be deleted at the end of the project.

The researcher: Liz Broadbent trained as a primary school teacher and has worked in many parts of England and the Netherlands. She has specialised in supporting young people with dyslexia and reading difficulties. She is particularly interested in finding additional ways of helping all young people learn to read.

If you have any questions before you decide contact me on liz.broadbent.14@ucl.ac.uk.

This project has been reviewed and approved by the UCL IOE Research Ethics Committee ref no Z6364106/2019/01/51

Supervisor: Professor Chloë Marshall, Department Psychology and Human Development, UCL Institute of Education

Data Protection & Freedom of Information Officer: data-protection@ucl.ac.uk

Appendix 4: Pupil consent form

I have been told about the research project. I understand what I will be asked to do.

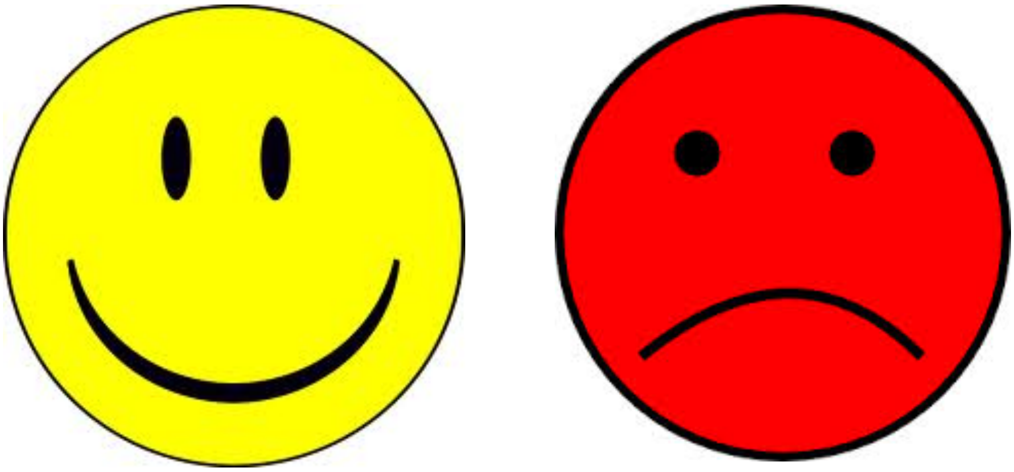
I would like to take part.

YES  NO 

Name………………………………………………………………

Class………………………………………………………………

Appendix 5: Continued consent, smiley face



These two smiley faces will be produced and laminated as a double sided disc

The pupils taking part will place the face uppermost indicating that they are happy to continue or wish to stop the process.

Appendix 6: Semi-structured interview

1.Do you like reading and read books for enjoyment (i.e. not your school reading book/ school text books)? *Tell me what sort of books you like best. When you go to the school library/ local library/ bookshop what sort of books do you choose? Do you ever read online/ tablet/ e-reader rather than printed books. If so what do you choose to read? Do you ever alter the font online/on e-reader etc.?* *NB: Q1 is to begin discourse and open conversation. Answers will only be reported if they illuminate answers to other questions.*
2. Did you prefer reading with [Arial*] or [OpenDyslexic*] or did the font make no difference to you? Can you explain your choice? *What did you like about your chosen font? Was there anything you did not like about the other font? Is there anything else you would like to say about the design of either font? Do you think either of the fonts helped you read better? Do you ever decide to read a book or not read a book because of the font it is printed in? Do you think the layout of a book or page (spacing) makes you read better?* *Fonts not named to participants, each font will be pointed to.

Appendix 7: Thematic coding (example)

1.Do you like reading and read books for enjoyment (i.e. not your school reading book/ school text books)? *Tell me what sort of books you like best. When you go to the school library/ local library/ bookshop what sort of books do you choose? Do you ever read online/ tablet/ e-reader rather than printed books. If so what do you choose to read? Do you ever alter the font online/on e-reader etc.?* I like reading especially Harry Potter. I like to choose books that look like a good story. I read on a tablet sometimes at home. I can make the writing change and make spaces. Yeah. I prefer a tablet because I can change it. *NB: Q1 is to begin discourse and open conversation. Answers will only be reported if they illuminate answers to other questions.*
2. Did you prefer reading with [Arial] or [OpenDyslexic[14]] or did the font make no difference to you? I prefer that one [OD] *Can you explain your choice?* That one [OD] you can look to see what line you are on much easier. The gaps between make it easy to tell. The letters look easier to read because they are thick at the bottom. The letters are easy to read because they are all different to each other. That one [A] are all the same and you could get muddled if you weren't thinking and read it wrong. *What did you like about your chosen font? Was there anything you did not like about the other font? Is there anything else you would like to say about the design of either font? Do you think either of the fonts helped you read better? Do you ever decide to read a book or not read a book because of the font it is printed in? Do you think the layout of a book or page (spacing) makes you read better?*

Participant EmJay, Control group, Arial default, SWRT 98

Font design: letters, thick base, boldness, uniformity

Clarity: letter shape

Spacing: lines

Reading performance: accuracy

Appendix 8: Research study findings report to schools:

The impact of OpenDyslexic font on the reading performance of young readers with dyslexia

OpenDyslexic default	Arial default	Arial expanded
The house was dark and no one was home. A burglar was in the bushes.	The house was dark and no one was home. A burglar was in the bushes.	The house was dark and no one was home. A burglar was in the bushes.

Every participant read passages printed in OpenDyslexic default font plus passages printed in either Arial default or Arial expanded. The passages were matched to their individual reading level. Each passage was scored for reading accuracy, reading speed and comprehension. The results were analysed to see if font design and/or font spacing made a difference to the scores.

Overall, there was a statistically significant improvement in reading accuracy and reading speed when the text was printed in OpenDyslexic font. This was generally the case for all readers whether they had reading difficulties or not. Individuals varied, but on average, reading with the text printed in OpenDyslexic font helped participants to make fewer mistakes and to read more quickly.

The different fonts did not affect the comprehension scores, so participants understood what they had read equally well when reading texts in OpenDyslexic or either version of Arial.

The difference in spacing of Arial default and Arial expanded made no difference to reading accuracy, reading speed or comprehension. This means it was the font design (the legibility and shape of the letters), not spacing, that was making a difference to reading accuracy and reading speed.

Unfortunately very few books are printed in OpenDyslexic but e-readers, e.g. kindle, kobo etc. include OpenDyslexic font, so changing font is straightforward.

OpenDyslexic can be downloaded on to tablets, laptops, PCs by visiting www.opendyslexic.org (put in 0.00 when asked to pay, the font is actually free to download). If required, text accessed on line can be copied and pasted into word and the font changed to OpenDyslexic to make reading easier. Documents can be written in word using the downloaded OpenDyslexic font.

A very big THANK YOU to all of the 78 pupils from the 7 primary schools who took part. Many thanks too to their parents and guardians for agreeing that they could take part and their teachers for allowing them time to come and work with me. A huge thank you to all of the Head Teachers and SENCos, without whose help and support this study could not have been completed.

Liz Broadbent

If you would like more details on the study or more information, email me on:
liz.broadbent.14@ucl.ac.uk

This project has been reviewed and approved by the UCL IOE Research Ethics Committee
ref no: Z6364106/2019/01/51.
Supervisor: Professor Chloë Marshall, Department Psychology and Human Development, UCL IoE
Data Protection & Freedom of Information Officer: data-protection@ucl.ac.uk