Carrasco Orozco, María

Can Virtual Reality improve Dyslexic English students' reading fluency and their emotional valence towards reading?

Master's Thesis in Education

FACULTY OF EDUCATION

Learning, Education and Technology, Master's Programme in Education

2020

University of Oulu
Faculty of Education
Can Virtual Reality improve Dyslexic English students' reading fluency and their emotional valence towards reading? (María Carrasco Orozco)
Master's Thesis in Education, 72 pages, 10 appendices
May 2020

The purpose of this master's thesis is to compare whether a text read in a virtual environment improves the reading of English students with dyslexia in terms of fluency compared with a text laid and read on a piece of paper. Furthermore, another interest of this study is to identify how the participants' emotional valence was aroused while reading.

This master's thesis is done with the help of Lyfta Oy, an EdTech learning company focused on 360º photos and VR learning environment. Moreover, the thesis design is based on a Lyfta's workshops, were one the participants read an entire passage in VR without difficulties being dyslexic. Two research questions are aimed to be answered during this research: (1) Is there a difference in fluency between reading in virtual reality (VR) and on a piece of paper? And (2) How positive and negative emotions were empathized while reading?

The study was carried out in the UK, were 23 Year 7, 8 and 9 students took part of the data collection. During this face, the participants were asked to read two short passages, one in VR and the other on a piece of paper, being video recorded and, they were asked to fill in two questionnaires about their emotions while reading both texts.

Afterwards, the number of errors, words read per minute and prosody were quantified based on the videos, to analyze the participants' fluency (which comprises three elements: (1) accuracy, (2) rate and (3) prosody) and to answer the first research question. The results suggested that there is not enough data to draw statistical difference between VR and paper.

To answer the second research question, the questionnaires' answers were analyzed. The results suggested that there is a statistical difference in terms of prosody and emotional valence between VR and paper.

This study could have some implications in school children having dyslexia, since it might boost positive emotional valence and hence boost their motivation to practice their reading skills. Moreover, educational companies might find a motivation to research more in depth in some aspects of this research and create educational products that can beneficiate dyslexic students' academic achievement. Also, this research could not only have an impact in dyslexic students, but in general education and other students, since the current master's thesis continues investigating and analyzing issues that are important in the school days and everyday life of students, such as the role of emotions in the classroom and how VR can affect their emotional valence.

Keywords: Dyslexia, virtual reality (VR), fluency, emotional valence

Contents

1 Introduction

The main aim of this master's thesis is to analyze if a text laid out in a VR learning environment makes a difference in the participant's reading, in terms of fluency compared to a text on a piece of paper for English students with dyslexia and moreover, to identify how positive and negative emotions were empathized and aroused while the participants were reading. The VR learning environment used is Awra Amba, created by the company Lyfta Oy (2016).

Dyslexia is a learning difficulty that affects how children (and adults) process information and it involves their literacy and language related skills (Reid, 2011; BDA, n.d). Dyslexia affects around 1 out of 10 people in the UK (NHS, 2018) and it is a lifelong learning difficulty, however, that does not mean that its effects cannot be mitigated. Nowadays, we can access to new and fast-growing technologies, that can provide us with new angles which we can look through a problem. The goal of this research is to use one of this fast-developing technologies (VR) and find out if it is beneficial for students with dyslexia in terms of their reading fluency and emotional arousal.

There can be found many researches about the uses of VR in intervention or assessment of different neurocognitive, neuropsychological or neurodevelopmental needs (such as dyslexia) (Attree, Turner and Cowell (2009), Kalyvioti and Mikropoulos (2014), Parsons et.al (2009), Rizzo et.al (2013) or Sigmundsson (2005), among others), but no research has been found about how specifically VR and dyslexia interrelate when it comes to the intervention of students' reading fluency. Therefore, with this research, the aim is to close that one gap and start raising questions so other researchers can continue investigating through the same path to possibly find out significant results that can affect positively to the fluency of children with dyslexia.

Furthermore, another important element of this master's thesis is the emotional aspect of reading and how VR could have an influence in how students' emotional valence changes. Emotional well-being is crucial, especially for dyslexic children (Reid, 2011). Emotions can facilitate and promote different learning strategies and self-regulative processes also; they are thought to have an influence on students' intrinsic motivation. Moreover, these processes and strategies can lead to an academic achievement improvement (Linnenbrink, 2007; Meyer and Turner, 2007; Pekrun, et.al., 2002). Because of that, this research also focuses on how emotions are perceived by the participants while reading, either in VR or on a piece of paper, and find

out if positive emotions are more aroused than negative, and if this fact can have a positive effect on students' wellbeing and educational achievement.

2 Theoretical Framework

The main theoretical concepts of this research are four: reading, dyslexia, VR, and emotional valence. The aim of this section is to facilitate the understanding of these main concepts and understand what has been already researched and where is the gap in the literature that needs to be filled and studied.

First, the theory of reading will be exposed. In this section, fluency will be discussed in length and several frameworks, which will be used later in the analysis section, will be presented. Then, dyslexia as a main concept will be explained and presented, moreover dyslexia in reading will be emphasized, since this study is mainly focused on what happens in the reading ability of children with dyslexia. The third concept that will be introduced is VR in learning and then more specifically how this technology has been used in the field of dyslexia. In another section, the literature in emotions in education will also be discussed and more specifically, emotions in students with dyslexia. Last, a review of earlier related studies will be exposed.

2.1 Theory of reading

Learning how to read is one of the most important tasks that children must face during their first years at school. Reading is essential to live and participate in our society, and whereas speaking is innate to human beings, writing and reading are not. It needs to be taught. (Miller, 2018).

When learning how to read in an alphabet script like English, "children have to integrate the meanings of words within phrases and sentences using knowledge of syntax and semantics" (Snowling, Goulandris & Stackhouse, 1997 p.86). It is a process that has several well-structured phases, according to Das (2009). It starts by reading very short words as if they were pictures, without making a relationship between the grapheme and the sound (car, cat, bat, sun…), as they progress, they start recognizing some words by sight and dividing them into parts (a-ni-mal, flo-wer…), but reading is not only decoding words, but also involves a certain level of comprehension of the context, because of that, these two entities are taught separately.

In summary, students go through these 4 stages:

- Symbolic stage, where lines are arbitrary for the children.
- Pictorial stage, where words are recognized as if they were pictures.

- Alphabetic stage, where the children realize that the words have specific sounds.
- Orthographic stage, where they start reading fluently. (Das, 2009)

But not everyone succeeds in this task, many children fail to learn to read at the end of their first academic year in elementary school. Many reasons have been hypothesized to explain this issue. Two commonly beliefs are, as reported by Das, a) the lack of exposure to reading material and b) a delay in the child's development and/or a specific deficit in intellectual skills that might affect their ability to read, write or talk, (Das, 2009). Also, it has been theorized, after the publication of the National Panel Reading report (2000), that the reason why some students don't succeed in reading might be because of a lack of reading fluency and effective instruction (Chard, Vaughn & Tyler, 2002; Kuhn & Stahl, 2000; Rasinski & Hoffman, 2003 as cited in Rasinski, 2004).

2.1.1 Fluency

Fluency refers to a person's language proficiency, as Housen and Kuiken (2009) point out, usually it is characterized by perceptions of ease, eloquence, and smoothness of the speech. Fluency reading comprises three important elements: accuracy, rate and prosody. (1) *Accuracy* refers to the ability to read a text recognizing and decoding words correctly, a good understanding of the alphabet, the ability to blend sounds together (Ehri & McCormick, 1998) and knowledge of many high-frequency words (Hudson, Lane & Pullen, 2005). All of these elements are required for a high accurate reading. (2) Reading *rate* encompasses, a good knowledge of high-frequency words that translates into a word-level automaticity (decoding words or recognizing them) and the speed in which the reader moves through a text (Hudson, Lane & Pullen, 2005). It is frequently quantified in terms of reading speed, either the time that takes for the reader to read a piece of text or the number of correct words per minute. (3) *Prosody* is a term related to the musicality and rhythmic aspect of the text. Several authors (Allington, 1983; Dowhower, 1991; Schreiber, 1980, 1991) explain that prosodic features are variations of intonation (pitch), syllable prominence (stressed words) and duration that helps the reader make sense of the text and provide expression to it. Also, in second language (L2) acquisition, we can find fluency conceptualized, according to Lambert and Kormos (2014), as "(i) *break-down fluency*, which relates to pausing behavior, (ii) *repair fluency*, which relates to the frequency of repetitions and self-corrections, and (iii) *speed fluency,* which relates to rate of delivery" (p.4).

For the sake of analyzing the data of this research, the first framework will be utilized since the language that the participants use during the data collection is English, their mother tongue, not an L2. There have been different approaches to measure how fluent a student is, according to the first framework, we can measure accuracy, rate and prosody.

Accuracy can be determined, according to Lambert and Kormos (2014) and Critchley (2005), by two similar approaches, (1) calculating the ratio of errors in a text by a unit, such as words, or (2), calculating the proportion of non-errors in a text, where the codes for the errors are substitutions (when a student changes a world for another), insertion (when the student adds a new word to the text) and omission (when a student omits a word in the text), told (when the teacher tells the word to the student) and appeals (when the student asks for help and the teacher tells the word). At the same time, a repetition of a word or an appeal that is not answered because the student can read the word independently are not coded as errors. Gickling and Thompson (1985), defined an instructional level of reading accuracy if the percentage of accurate reading words was between 93% and 97%. Other authors agree that if the percentage falls below 93% the reading accuracy level needs to be reinforced (Beck, Burns & Lau, 2009; Burns, 2007; Burns, Hodgson, Parker, & Fremont, 2011).

Rate is most of the time quantified as the reading speed. As it was already exposed, according to Hudson, Lane and Pullen (2015), it can be done calculating the amount of correct words read per minute or the time it takes to the reader to complete the full passage. Finally, prosody cannot be measured objectively since it cannot be quantified (Hudson, Lane & Pullen, 2005). It needs to be measured through an observation of an oral practice. To measure students' prosody, we can follow the following rubric created by Zutell & Rasinski (1991) and adapted by Rasinski (2004, p.3,4).

A. Expression and Volume

1. *Reads words as if simply to get them out. Little sense of trying to make text sound like natural language. Tends to read in a quiet voice.*
2. *Begins to use voice to make text sound like natural language in some in areas of the text but not in others. Focus remains largely on pronouncing the words. Still reads in a quiet voice.*
3. *Makes text sound like natural language throughout the better part of the passage. Occasionally slips to expressionless reading. Voice volume is generally appropriate throughout the text.*

4. *Reads with good expression and enthusiasm throughout the text. Varies expression of volume to match his or her interpretation of the passage.*

B. Phrasing

1. *Read in monotone with little sense of phrase boundaries; frequently reads word-by-word.*
2. *Frequently reads in two- and three-word phrases, giving the impression of choppy reading; improper stress and intonation fail to mark ends of sentences and clauses.*
3. *Reads with a mixture of run-ons, mid-sentence pauses for breath, and some choppiness; reasonable stress and intonation.*
4. *Generally, reads with good phrasing, mostly in clause and sentence units, with adequate attention to expression.*

C. Smoothness

1. *Makes frequent extended pauses, hesitations, false starts, sound-outs repetition, and/or multiple attempts.*
2. *Experiences several "rough spots" in text where extended pauses or hesitations are more frequent and disruptive.*
3. *Occasionally breaks smooth rhythm because of difficulties with specific words and/or structures.*
4. *Generally, reads smoothly with some breaks, but resolves word and stricture difficulties quickly, usually through self-correction.*

D. Pace

1. *Reads slowly and laboriously.*
2. *Reads moderately slowly.*
3. *Reads with an uneven mixture of fast and slow pace.*
4. *Consistently reads at conversational pace; appropriate rate throughout reading.*

The authors define an instructional level of prosody if the scores were higher than 8. If it is lower, it indicates a need for further reading reinforcement.

2.2 Dyslexia

Dyslexia is perhaps the most common learning difficulty affecting children and it occurs in all languages (Shaywitz, Morris and Shaywitz, 2008). It is estimated that around 15 to 20 percent of the population show some symptoms of dyslexia, which includes slow and inaccurate reading or poor spelling (International Dyslexia Association, 2017).

The first description of reading alterations was made my J. Kerr in 1896 and it wasn't until 1917 that the term dyslexia was proposed by Hinshel Wodd. The word dyslexia come from the Greek prefix *dys*, which means "poor" or "inadequate" and *lexis*, which means "words" or "language". Literally it means "inadequate language" or "problem with words" (Cox, 1985). It has been difficult to agree on a definition for this learning difficulty, since it was complex to determine how and why it happened, but two definitions are highlighted.

According to The International Dyslexia Association (2017), "Dyslexia is a language-based learning disability. Dyslexia refers to a cluster of symptoms, which result in people having difficulties with specific language skills, particularly reading. Students with dyslexia usually experience difficulties with other language skills such as spelling, writing, and pronouncing words. Dyslexia affects individuals throughout their lives; however, its impact can change at different stages in a person's life. (…)"

The British Dyslexia Association (2007) defines it as: "Dyslexia is a specific learning difficulty that mainly affects the development of literacy and language related skills. It is likely to be present at birth and to be life-long in its effects. It is characterized by difficulties with phonological processing, rapid naming, working memory, processing speed, and the automatic development of skills that may not match up to an individual's other cognitive abilities. It tends to be resistant to conventional teaching methods, but its effect can be mitigated by appropriately specific intervention, including the application of information technology and supportive counseling."

Diagnosing dyslexia is a difficult task due to the nonexistence of a clear cut-off between a dyslexic and a non-dyslexic reader (Snowlig & Hulme, 2012). There are many hypotheses that try to explain the causes of dyslexia, such as, the rapid auditory processing theory (Tallal, 1980, 2000; Tallal, et al, 1993); the visual theory (Livingstone et al, 1991; Lovegrove et al, 1980); the cerebellar theory (Nicolson and Fawcett 1990; Nicolson et al, 2001); and the magnocellular theory (Galaburda et al, 1994; Livingstone et al, 1991; Stein 2003; Stein and Walsh, 1997).

But the predominant cause hypothesis of dyslexia is thought to be the deficit of phonological skills that affect the acquisition of literacy skills, such as reading, spelling or writing (eg. Vellutino, Fletcher, Snowling & Scanlon, 2004). Some of the predictors of dyslexia within alphabetic orthographies are phoneme awareness, together with letter knowledge and rapid automatized naming (RAN), sometimes referred to as "code related skills" (Caravolas et al., 2012; Moll et al., 2014; Ziegler et al., 2010). However, according to Snowling and Melby-Lervag (2016), those predictors are not enough to explain dyslexia. There are other factors to take into account when diagnosing it, such as family risk (FR); indeed, FR is an strong factor, since about 65 percent of children who suffered from dyslexia reported having a parent with the same disorder (Scarborough, 1990). Several studies carried out with a 4-fold comparison to a control group reported differences in speech and tone of 6 months old babies (Leppänen et al., 1999; 2002; Guttorm et al., 2001), differences in phoneme categorization also in 6 year old's brains (Richardson et al., 2003) and poor cognitive skills that are known to predict their future ability to process letters, rapid automatized naming or phonological processing (Lyytinen, P. & Lyytinen, H., 2004, Lyytinen, P. et al., 2005). Hence, the inter-relationships between a phonological aspect and other possible factors needs further investigation.

There are different kinds of symptoms and manifestations of dyslexia in children, such as manifestations in their reading or writing, but there's no relevant empirical data that agrees in behaviorism symptoms. However, "empirical literature suggests that children, adolescents and adults with dyslexia are at risk of low self-esteem" (Burden, 2008 in Nalavany, Carawan & Brown, 2011, p.1), but it is thought to happen not because of the reading difficulty itself, but because of the context that surrounds the child, since evidence show that dyslexic children receive less help than non-dyslexic students in mainstream schools, making it more difficult for them to keep up and consequently creating extra stress (Riddick, 2006). That is why, in this research we will focus not only in the reading quality but also in the emotions that reading aloud brings to the students.

There are three classifications of children with developmental dyslexia (disorder of congenital or hereditary origin, different from acquired dyslexia, which is a disorder caused by a brain injury (Sprenger-Charolles, Colé & Serniclaes, 2006).), based on diagnostic reading-spelling patterns, according to Border (2008), which are: dysphonetic dyslexia, dyseidetic dyslexia and, mixed dysphonetic-dyseidetic dyslexia.

The first one refers to children whose reading-spelling pattern shows a deficit in the integration of grapheme-phoneme (symbol-sound). Children within this classification have a limited sight vocabulary, when they see a word that is not in their vocabulary, they are unable to decipher it because of the lack of word-analysis skills. Regarding a spelling pattern, they try to spell a word by sight and not "by ear". The second one refers to children whose reading-spelling pattern shows deficit "in the ability to perceive letters and whole words as configurations" (Border, 2008, p. 667). The last one refers to children whose reading-spelling pattern shows a deficit in the two already explained classifications.

2.3 Virtual Reality in Learning

Virtual Reality "refers to a whole stimulated reality, which is built with computer systems by using digital formats. Building and visualizing this alternative reality requires hardware and software powerful enough to create a realistic immersive experience (e.g. VR helmets or dedicated glasses and 3D software)" (M. Gutiérrez, et.al., p.473). Rosenblum and Cross (1997) defined already three aspects that make part of any VR environment and that is still current nowadays. Those aspects are immersion, interaction and visual realism; therefore, virtual glasses, surround sound, etc. is essential to simulate a realistic environment (Wu, et al., 2015).

This technology was first used in the mid-80s (Franklin Institute, 2018) and since then has evolved in a very fast pace, been used in many different areas thanks to its adaptability and flexibility (LaValle, 2017). VR has the affordance to induce a specific behavior in an organism by utilizing artificial sensory stimulation, according to LaValle (2017), allowing users a very similar behavior and reaction from reality (Sherman & Craig, 2003), which can help researchers recreate some situations that might be difficult to do in reality.

This technology has been proven effective in different domains, for example, in museums communicating cultural content to users with different cultural backgrounds (Carrozzino & Bergamasco, 2010), or in classrooms, improving a learning environment (LE) by allowing social interaction (Bailenson, Yee, et al., 2008) or allowing users to focus better on some tasks and providing better memory recall ability (Krokos et al. 2018). Moreover, teachers with the help of VR, have the ability of travelling with their students to other historical eras, exploring the space or even the interior of a human-being, also another affordance of VR is the increase of the engagement of even increase the educator's engagement as well (Peachey, et al., 2010).

However, there are some limitations and need for more research, since “it still remains unclear to some extent where the real benefits and limitations of using virtual worlds as knowledge transfer and learning environments are compared to more traditional methods” (ViWo 2008, n.p).

As Fowler, (2014) states, the use of VR holds great promise when creating LE, but also holds many challenges, such as understanding the pedagogic behind this LE that should give information about the design and the use of these VR environments. Mikropoulos and Natsis (2011) reviewed over 50 papers in a span of 10 years. One of their observations was that very few of the reviewed studies about VR in education contained clear information about the pedagogical model used in their LE and VR environments, as they stated, “all the other reviewed articles do not refer explicitly to a learning theory” (Mikropoulos and Natsis, 2011, p.775).

Other limitations not related specifically to education are the feeling of motion sickness, dizziness, nausea or headache while in an immersive environment using a VR headset (which is currently needed to enjoy VR) (Feng, 2018). Head-mounted displays, such as VR, differ from other vehicles in terms of inertia displacement, since VR does not move relative to the Earth (Kennedy, et.al, 2010). Hence, as stated by Feng, “a thoughtful design of IVR (Immersive Virtual Reality) (...) is necessary to minimize these side effects.”

The relationship between dyslexia and virtual environments, such as VR, is still not well defined, but few attempts have been made in some studies, such as the one from Kalyvioti and Mikropoulos (2014), presenting benefits and orientations for families and learning professionals (Cidrim & Madeiro, 2017).

2.4 Emotions in Learning

Emotions are a crucial part of learning, affecting everyone in the schooling process. Understanding the nature of them in the school context is vital for a healthy development of the students’ emotional wellbeing. (Schutz, Quijada, de Vrie & Lynde, 2010). Since there is no consensual conceptualization and operationalization of what to be studied (Scherer, 2005) they are a term complex to define. A possible definition could be a multifaceted phenomena that involves some sets of coordinated psychological processes, such as affective, cognitive, physiological, motivational and expressive components (Shuman & Scherer, 2014). According

to Schnall, “an emotion typically follows a specific eliciting stimulus or event and is intense but limited in duration” (2010, p.59).

Emotions and moods are often compared since they share similar characteristics and components, but moods are not as intense as emotion and they lack a specific stimulus or referent. They can be regarded as low intensity emotions (Prekrun, 2006). In this master’s thesis, it is referred as emotions since the participants will be given a specific stimulus (the reading task) during the data collection.

Schnall (2010), also affirm that emotions can be classified in two dimensions: valence and arousal. Valence can be positive (pleasant) states, such as happiness or enjoyment and negative (unpleasant) states, for example anger or boredom. Arousal entails activation and deactivation states. Therefore, by classifying these emotional states as positive/negative, and activating/deactivating, a 2x2 taxonomy can be created: (a) Positive activating: e.g. enjoyment, pride, excitement, etc. (b) Positive deactivating: e.g. relief, relaxation, etc. (c) Negative activating: e.g. anger, anxiety or frustration, etc. (d) Negative deactivating: e.g. sadness, exhaustion, etc. Moreover, in addition to valence and arousal, emotions can be classified as approach (when an emotion makes a person more likely to confront the source of such emotion, like anger) and avoidance (when a person is more likely to withdraw from other people, such as sadness). They can also be grouped depending on their object focus (Pekrum, 2006), and it determines if the emotion that the student/teacher is having is happening because of an academic task or not or in other words, if the specific emotion is dependent of an academic task or not.

There is a consensus that dyslexia has a deep effect on how the students’ educational experience is and their ability to master literacy skills, but it is still little what we know about the emotional effects of dyslexia in children and how they affect their day to day life and academic life (Palti, 2002). However, we know that emotional well-being is crucial to effective learning and especially for children with dyslexia (Reid, 2011). Research (Hales, 2009; Miles, 2004), suggest that due to a failure to acquire the necessary literacy skills to succeed at school, children with dyslexia can show attention and behavioral problems and this could lead to greater emotional problems if not acknowledge in time. Moreover, stress is a common symptom of this learning difficulty if, as it has been mentioned before, school failure prolonged in time and is not dealt with it in time. Studies (Riddick, Farmer, Serling,1997) comparing dyslexic and non-dyslexic children show that dyslexic ones are more likely to suffer anxiety than their peers. This stress

and anxiety levels (Mueller, 1976), can affect the student's ability to study, resulting in poor memory and problem-solving skills.

Another problem found in children with dyslexia is that most of them have difficulties recognizing their own success and appraising their own abilities. Studies suggest that the role of self-esteem and perception of one's abilities is crucial for dyslexic children and adults (Reid, 2011) as well as self-awareness, that is considered a factor in successful for an individual with a learning difficulty (Goldberg, Higgins, Raskind & Herman, 2003). Motivation is another key word when it comes to the emotional development of children with dyslexia and also many other students with learning needs. As Reid mentions, "many children with dyslexia find it difficult to become and stay motivated with some tasks (2011, p.141).

2.5 Review of related studies

There are few studies relating dyslexia, VR and specifically about how VR would influence students' fluency and emotional valence. But there is plenty relevant literature for this thesis, that motivates research in the use of VR in children with specific learning difficulties, such as dyslexia. The following paragraphs will be divided into overarching themes that relate to this master thesis. These themes are VR and learning, VR and learning needs, where dyslexia is part of, reading and dyslexia and experienced emotions in reading among people with dyslexia.

Research suggest that the use of VR in education results in a fomenting and improvment of motivation, engagement and some other skills. Cidrim and Madeiro (2017), researched the use of information and communication technologies (ICT), which includes VR, in education and their findings suggest that "the use of ICT in education seems to provide more interactive experiences, which can motivate children, attenuating the impacts of their own difficulties in the daily practice of reading and writing" (p.105). Also, following the same line, Schneider (2019) studied if there were any difference in students' engagement levels after using immersive VR and the results supported her hypothesis with a statistically significant correlation. Moreover, Mikropoulos and Natsis (2008) also posited in their study similar results. They found out that teachers and students show positive emotions and attitude towards VR, however they also mention that technological affordances like immersion and individual factors needed to be further studied in relation to learning outcomes. But research does not only suggest improvement in students' engagement or motivation, in a study conducted by Pellas, et al. (2017), it was found that virtual worlds (VR) had a potential for collaborative problem solving

and for the use of high order skills, which are often used also to keep learners' engagement. Nevertheless, as Mikropoulos and Natsis mentioned, there is a need for more research in the uses of VR in terms of its outcomes in education. Fact that is shared also by Freina and Ott (2015), who agree on the limited research on the effects of VR on children and its application to the classroom.

The use of VR on students with different learning needs have been divided mostly in assessment, intervention or training (Kalyvioti & Mikropoulos (2014). In a literature review carried out by the same authors, the effectiveness of VR as a clinical tool for increasing and improving awareness of dyslexia was studied. Their study focused on four domains of study: assessment, intervention, training and informative/awareness and their results showed that "almost all studies showed positive results (…) virtual environments have been useful applied for clinical assessment purposes, cognitive neurorehabilitation and in clinical psychology domains" (p. 142). Moreover, their findings and other studies, such as Rizzo et.al (2013) and Parsons et.al (2009) agree that VR applications can be used as a clinical tool for neurocognitive, neuropsychological or neurodevelopmental (such as dyslexia) impairments.

Different types of studies have been carried out about VR and dyslexia. The following are some of those. Sigmundsson (2005), assessed and compared dyslexic drivers to non-dyslexic driver's response time when driving a car, showing with his results that dyslexic drivers performed poorer than non-dyslexic drivers and arguing that this results support the hypothesis previously proposed by (Stein & Walsh, 1997) that many of the problems of dyslexic students originate from the difficulty of processing information at a fast pace. Another example of the use of VR on dyslexic students can be seen in the study of Attree, Turner and Cowell (2009), who constructed a virtual bungalow to identify visuospatial skills in dyslexic students. Their results showed that dyslexic students performed less well than children without dyslexia, however dyslexic children performed better on a real-spatial test than non-dyslexic students. This fact leads to the proposal that "dyslexia could possibly be associated with superior performance on assessments that reflect real-world tasks". Kalyvioti and Mikropoulos (2012), obtained good results in the creation of six virtual environments to study undergraduate student's memory skills, that had been diagnosed with dyslexia.

Another study regarding VR as a tool, in this case for attention training in dyslexic students (Pedroli, et al., 2017), found positive results in their participants' attention, and while their study results did not elicit an immediate effect on dyslexic students' reading performance, it suggested

that with more time there might be a significant improvement, since their results showed a significant decrease in the reading time of low frequency words. Lastly, should also be mentioned the study conducted by Passing (2011), to raise dyslexia awareness. The author used ten immersive virtual environments to help parents understand their dyslexic children and therefore, increase their awareness.

Moreover, other researches about dyslexic learners' reading using different ICT supports, were found relevant for this thesis. For example, in a study carried out by Schneps, et al. (2013) about the use of e-readers with children with dyslexia showed positive results in the participants' reading ability because of the e-reader affordances, such as font size and type, which is the same as can be done using a virtual reality environment, such as the one that will be used for this thesis. Another study (Sakamat, et.al, 2017) suggests that multimedia and digital storytelling support dyslexic learners which are found to be more motivated using multimedia programs. Also, Sakamat and colleagues agree that this method helps reducing the difficulties dyslexic students encounter when having to recognize familiar words. This last study is quite interesting for the current thesis, since the environment, Awra Amba, used from the host company Lyfta is a multimedia (containing VR, 360 images and video) storytelling platform and maybe its affordances are similar to the ones mentioned before for dyslexic students. Continuing with the articles using different technology supports, other than VR, for the study of dyslexic learners' reading, eye tracking has started to be used as a screening method for reading; Benfatto and others (2016), used this technology to assess and demonstrate that "it can produce individual-level predictions with high sensitivity and specificity in less than one minute of tracking time" (p. 13), with positive results that ascertain their hypothesis.

Last, it is important to review studies regarding emotions and motivation. Positive emotions and their influence in students's academic achievement is an important part of this research, since the second research question focuses mainly on them. Positive emotions play an important part in individuals with a learning need, especially on dyslexic individuals, since study shows how depression and anxiety are often reported in them (Nalavani, et al., 2017). Moreover, findings (Kalka & Lockiewicz, 2017) present that students with dyslexia felt less happy than non-dyslexic learners and showed a lower level of positive emotions. Therefore, the importance of a positive emotional valence. It has been reported (Abedi, et al., 2019) that they help improving the image that an individual has on oneself and some of them, such as happiness, improves well-being and "plays a greater role in improving this construct", (p. 11).

Motivation is also an important part of this study as well as how the emotional valence of the participants is highlighted when reading, either in VR or on a piece of paper. In a study from Huang and Liaw (2011), findings suggested that VR contributed to enhance students' motivation an 18% and the authors concluded, supporting the idea that VR can be an antecedent to motivation (Lee, et. al., 2010), that "VR technology can increase interactivity and immersion of course contents, thus providing and immersive, engaging environments which can stimulate learners' motivation and curiosity" (p.302).

3 Aim and Research Questions

This master's thesis project came into existence thanks to an EdTech company named Lyfta Oy (https://www.lyfta.com/).

Last year, Lyfta visited a school in North West London, while running a VR based workshop with a group of year 8 children. There were 6 children in the group, using one VR headset among them. The workshop conductor asked the children to describe what they could see and hear while wearing a headset to keep the whole group engaged throughout the session. The second child to read, we will call him Charles due to privacy arrangements, explored the weaving factory in a small Ethiopian village called Awra Amba, using Lyfta's VR environment.

The conductor asked Charles to open some of the available multimedia content and read out the text for everyone to hear. What the conductor didn't know, was that Charles had dyslexia. Charles hesitated for a moment, but then went on to read the short passage in front of his peers, which he delivered correctly. When he took the headset off, he told the conductor that it had been the first time he had read aloud without making a mistake.

After hearing this story, I decided to delve more into the issue by conducting a formal study about the effects of VR learning environments on reading for children with dyslexia. The aim of this master's thesis is to find out if reading using a VR environment and with all its affordances, such immersiveness, will improve how English dyslexic students read in terms of fluency and also to see how emotions play a part in the students' reading.

In order to ascertain our hypothesis, two research questions will be answered, which are as follows:

RQ1 Is there a difference in fluency between reading in VR and on a piece of paper?

RQ2 How positive and negative emotions were empathized during reading?

To be able to collect the data, Lyfta and I visited the UK in two occasions. The first one was to attend a Special Education Needs conference (NAHT, 2018), where we expected to meet school principals who might be interested in our study. After the conference we sent emails to all these school principals and tried to recruit other schools and teachers to help us with our study. Furthermore, a website was created (https://www.lyfta.com/dyslexiastudy) (see appendix 1), where the interested schools could check the details of the study and reach us. Our goal was to

have around 50 participants. The second occasion we travelled to the UK was to collect the data.

4 Methodology

The present study is a quantitative research based on a series of video recordings and questionnaires. In this chapter, the data collection design will be explained, as well as the instruments used in the study. Continuing the discussion, the participant pool, consisting of year 6, 7 and 8 children, will be described and finally, the data collection procedure will be presented.

4.1 Participants and Context

The sample used for this study consisted of 23 year 7, 8 and 9 students, from which 2 of them were excluded due to anomalies while carrying out the data collection. The gender of the students was equally distributed: Male, N=11 (52%) Female, N=10 (48%).

The students' ages ranged from year 7 (6th grade), age between 11 to 12 (N=13), year 8 (7th grade), age between 12 to 13 (N=6) and year 9 (8th grade), age between 13 to 14 (N=4), and 11 out of the 21 valid participants had used VR before.

The data was collected from two different schools within the UK; Ysgol Abersychan School in Pontypool, Wales and Rossendale School in Ramsbottom, England. The data collection took a week and each student session lasted between 30 to 60 minutes.

4.2 Data Collection Design

To carry out the data collection it was decided to use a video-based observation method to measure children's fluency, as well as questionnaires to analyze their emotional arousal. With the idea of replicating the previously explained workshop, two texts from Lyfta's Awra Amba storyworld were selected. Both with similar difficulty levels, in terms of the vocabulary used. We prepared the same text content in both, printed media and VR, using slightly different approaches:

(1) The paper versions had one image at the top or bottom of the page, a layout very similar to standard textbooks used in schools and contained 5 chapters of 1 paragraph per chapter. The font used was *Univers Lt St*d, roman for body and bold for titles (body: size 10, leading 12 and title: size 14, leading 16). An example can be seen in appendix 2.

(2) The VR version was slightly more complex in terms of design. The font used was *Univers Lt Std,* roman for body and bold for titles. We planned for the participants to enter a VR environment visually themed for the topic, and for them to enter the correct space inside (there were two texts versions). In there, slides could be seen and read – each consisting of one chapter (one paragraph just like in the printed version). An example can be seen in appendix 3.

To ensure the reliability and validity of the research, *Univers Lt Std* font was chosen, with the above specifications, because it was desired that both texts, despite being laid out in different environments, were as similar as possible to textbooks that our participants could find at their schools, since it was intended to compare a text in VR against the texts children have to read at schools. To be able to replicate them, we referenced and inspired our texts styles and similarities on the Geography book Horizons 1: Student Book, a book that is used in schools within the UK.

Two different texts were selected, which will be referred to as A and B and both texts had their equivalent versions printed and in VR. Considering that the participants were going to be recorded and they had to read both texts, one after the other, meaning that they were going to be reading for a while, it was possible that a learning curve happened since they could get used to the camera and to reading as time passed and so, improve their reading skills which would skew the data. Because of that and with the intention to avoid this learning curve, four different reading approaches were established regarding the reading order of the texts (A or B) and how and when the participants would read (VR or Paper).

Some participants would start reading one text (A or B) aloud (since the idea is to test their reading accuracy aloud), on paper, and then proceed to read the remaining text in VR, and for the other participants it would be the other way around, first text A or B in VR and then the remaining text on paper. Thus, there are 4 different approaches:

- Paper, text A then VR, text B
- Paper, text B then VR, text A
- VR, text A then Paper, text B
- VR, text B then Paper, text A

As mentioned, fluency consists of accuracy, rate and prosody. In order to collect data of all the categories without disrupting or interfering with the experiment nor the participants and

because “video data might improve ecological validity, since the video data gives more complete (and visual) information about the real environment rather than traditional observers’ observation notes” (Asan & Montague, 2015, p.nf) it was decided to proceed with a video-based observation method.

The experiment setup was as follows: there was a table and a chair in a classroom where each participant sat down to read the texts. In front of every child there was a camera recording their reading. In order to prevent possible stress caused by the participants being recorded, the camera was set up and turned on before every participant arrived. Only after the participant left was the camera turned off.

The video recordings were edited in smaller chunks for a better analysis. After that, they were seen, and the aforementioned categories quantified to later on be analyzed with a quantitative method using different statistical test.

As for the second research question, two questionnaires were prepared to assess participants’ emotions after they read on paper and in VR. They were designed combining an image-based Likert questionnaire and a Likert mood scale (Lyanne, et al. 2016 and Reynolds-Keefer, 2011).

Since the data would be collected working with dyslexic children, it was desired that they could answer the questionnaires without the bias of having to read different options to answer. There is evidence that the use of emojis or emoticons as a method in questionnaires for research disambiguate the communicative intention behind the message (Kaye, 2016) and also, serve important verbal and nonverbal functions in communication (Yuasa, 2011).

Moreover, according to Pink (2003) in Fane (2016), “emerging technologies have the potential to produce “new, innovative, reflexive, and theoretically informed research” (p.191), through their ability to accommodate different audiences and purposes” (p.3, 4). Also, in a research carried out by Marsh (2005), it is discussed that the increased focus on multimedia and technology supports the use of emoji as a research method to engage students and help them understand better.

Because of the above reasons, it was agreed that the use of emoji/emoticons would in this case help with the understanding of the questionnaires, hence it was decided to use an image-based Likert questionnaire and Likert mood scale.

Table 1: Summary of data collection design

	Data collection	Analysis	
RQ1	- Video recordings (reading on paper and in VR)	- Categorization and quantification (types/number of errors, time per words read, prosody.)	- Paired-Sample T-test - Wilcoxon Signed-Rank test
RQ2	- Image-based-Likert questionnaire - Likert mood scale	- Wilcoxon Signed-Rank test	

The questionnaires used are presented below in figure 1 and figure 2.

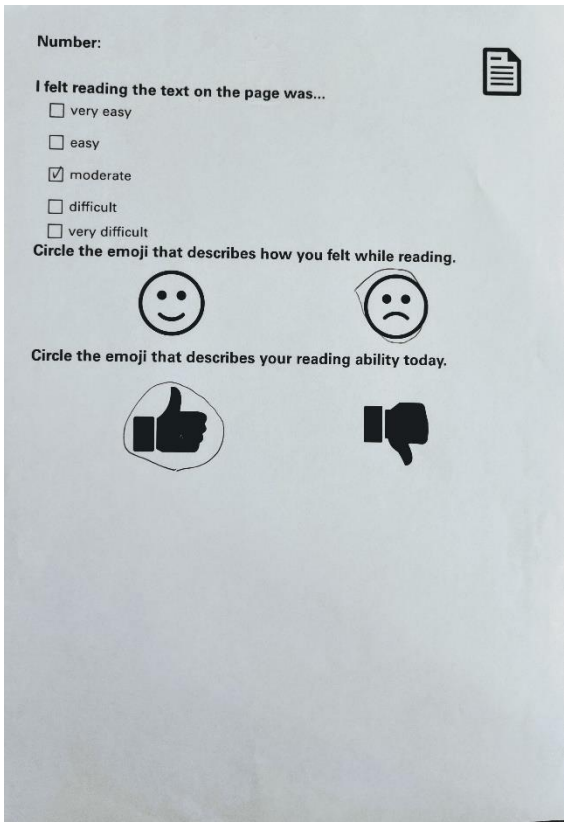


Figure 1: Questionnaire after reading on paper

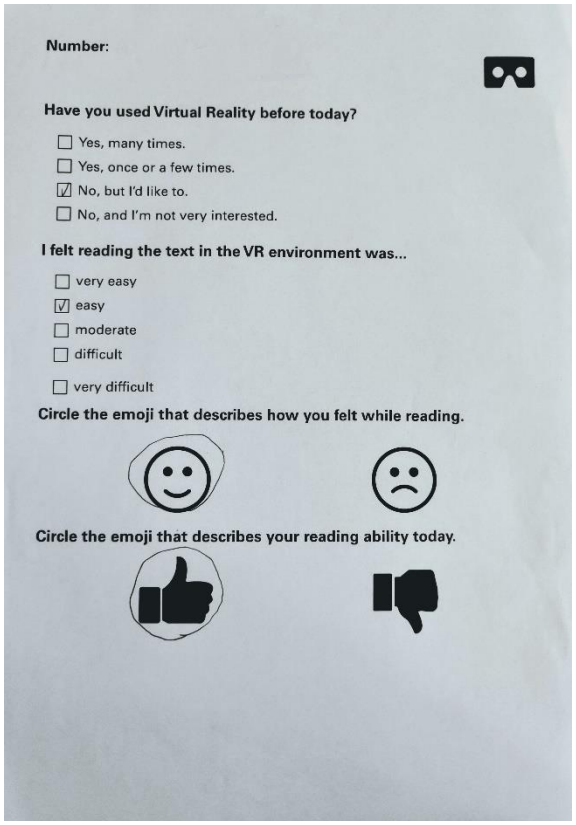


Figure 2: Questionnaire after reading in VR

4.3 Data Collection Procedure

The participants were asked to read a total of two passages and fill in a questionnaire after reading each passage (a total of two questionnaires per student). They read one of the texts in a piece of paper, simulating a textbook and the other piece of text in a VR environment, with some of its affordances, such as sound and immersiveness. Table 2 shows how the reading order was decided and varied, as well as the amount students who read in that specific order and the students' gender.

Table 2: Reading order during the data collection with number of participants and their gender

VR – text A	3 females
Paper – text B	2 males
	1 male outlier
VR – text B	3 females
Paper – text A	2 males

Paper – text A	2 females
VR – text B	3 males
	1 female outlier
Paper – text B	2 females
VR – text A	3 males

4.4 Data Analysis

The research method used in this study is a quantitative method. "Quantitative research involves the collection of data so that information can be quantified and subjected to statistical treatment in order to support or refute alternate knowledge claims" (Creswell, 2003, p. 153). The reason a quantitative method was chosen is because fluency is mostly measured quantitatively and the frameworks and methods found to measure it and its subcategories (accuracy, rate and prosody) were quantitative. Moreover, there was a time constraint when collecting the data collection, and if a qualitative research had been chosen, a longer time would have been needed to carry out a longitudinal study (Thomson & McLeod, 2015) for robust results. Therefore, since a quantitative method was used, the measured variables fell into different quantitative data types, which can be seen in figure 3.

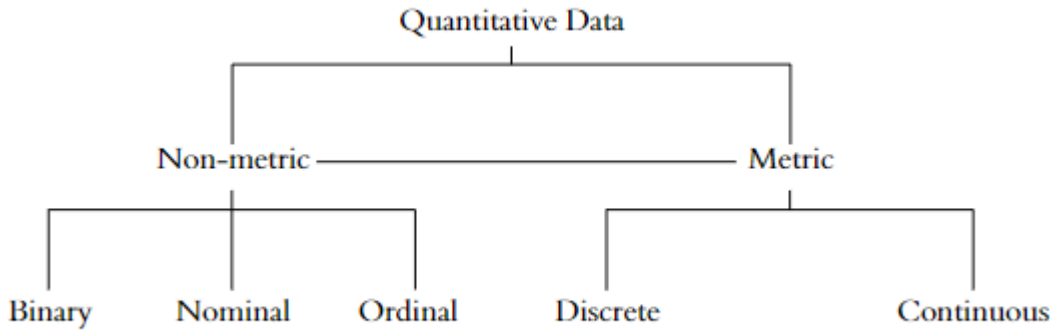


Figure 3: Classification of Data Types (Singh, 2007, p. 123)

Part of the data was collected using a video recording method, which was coded using the software NVivo 11 and once the data was divided into categories, a quantitative method was implemented using the software IBM SPSS Statistics 25 (SPSS) in order to answer the first

research question, which was: is there a difference in fluency between reading in VR and on a piece of paper? The resto of the data and to be able to answer the second research question (how positive and negative emotions were empathized while reading?), the same software to analyze quantitative data was used (SPSS).

4.4.1 Coding the videos for RQ1

Fluency is a category umbrella that encompasses three subcategories: Accuracy, Rate and Prosody. For accuracy, errors were quantified. Using NVIVO, the videos were watched and whenever an error was made, it was divided into three different codes which were substitutions, insertions and omissions. Examples of those coding schemes can be seen in table 3, where the error committed by the participant is bolded for a better understanding of the type of error.

Table 3: Example of accuracy data coding

Original text	Read text by participant	Code
"(…), **there** was a need to open a supermarket"	"(…), **where** was a need to open a supermarket"	Substitution
"(…) to other farmers by the Government during a land reform"	"(…) to other farmers by the **other** Government during a land reform"	Insertion
"Our mill attracts thousands of costumers"	"Our mill thousands of costumers"	Omission

Rate was quantified by calculating how many words the participants read per minute. It was done by dividing the total amount of words per text between the number of minutes read. An example of how rate was quantified in participants 1 and 2 can be seen in table 4.

Table 4: Example of rate data quantification

Participant	Text	Reading environment	Words per minute
1	A	VR	42.60869565
1	B	Paper	54.52471483
2	B	VR	68.28571429
2	A	Paper	67.35483871

Prosody was coded following a framework created by Rasinski (2004, p.3,4) and shown in the literature review section of this thesis. Since the reading ability of the participants was based on a rubric and therefore is subjective, an inter-rater reliability test, specifically a Cohen's Kappa test, was done to ensure the reliability of the results, where an external rater coded 100% of the prosody data. According to Fleiss agreement guidelines (1981), the test results show that there was an excellent agreement (the mean of the tests was k=.81) between the two raters in each of the categories both in paper and VR. An example of how prosody was coded can be seen in table 5.

Table 5: Example of prosody data coding

Participant	Text	Reading environment	Expression and Volume	Phrasing	Smoothness	Pace
1	A	VR	3*	2	3	3
1	B	Paper	3	2	2	3
2	B	VR	3	2	3	3
2	A	Paper	3	4	4	3

*The correspondence of the numbers can be seen in the literature review section (prosody), in the rubric created by Rasinski (2004, p.3,4).

4.4.2 Quantitative analysis for the RQ1

Accuracy is analyzed using the software SPSS. For the first analysis all the errors made by the participants were considered, ignoring the categorization of them, and were compared between paper and VR. Since the lengths of the texts varied, the variable to analyze was error rate (e_r) (instead of number of errors), calculated in equation 1, where e_r is error rate, e_t is total errors and w means words.

$$e_r = e_t / w \tag{1}$$

Since it is desired to know if the difference between paper and VR and the variable is continuous, there was an inclination to use a T-Test and since the variables were measured twice for the same participant and the datasets (Paper, VR) are not independent, a Paired-Sample T-Test was the chosen test.

A Paired-Sample T-Test is used to analyze if the means of each dataset are significantly different and it should be used if our data fits the test's assumptions, which were as follows:

- *Independent observations or, more precisely, independent and identically distributed variables.*
- *The difference scores between the two variables must be normally distributed in our population.*

The first assumption was met since we had different participants. As for the second assumption, to check if our data is normally distributed, a Shapiro-Wilk Test of Normality was used, because the sample size is small (< 50 samples). After doing the test, it can assure that the data fits a normal distribution, since the p-value of the test was greater than 0.05 (Sig.=.560) (see appendix 4).

Afterwards, a more thorough analysis was carried out now focusing on each of the categories. Again, we wanted to know the difference between paper and VR and now more specifically between insertions in paper and insertions in VR, substitutions in paper versus substitutions in VR and the differences between omissions in paper and omissions in VR. Following the same process as e_r , equation 1 was used on the errors of each type, resulting in e_r^i = error rate for insertions, e_r^s = error rate for substitutions and, e_r^o = error rate for omissions. Since the dataset has the same characteristics (the variables are continuous, and the data is paired) it was decided to use a Paired-Sample T-Test.

The variables fit the assumptions, which means that they are independent and identically distributed and also, according to the Shapiro-Wilk Test of Normality, they meet the normality requirements (insertions Sig.=.072; substitutions Sig.=.118 and omissions Sig.=.177) (see appendix 5).

Regarding rate it was analyzed using the software SPSS. Again, since the texts have different lengths, words per minute (W_m) were calculated in equation 2 (where W_m means words per minute, w is number of words and t_m means time in minutes.

$$W_m = w/t_m$$

(2)

A Paired-Sample T-test was used to analyze rate since it is desired to know if the difference between paper and VR the variables are measured twice (same as with accuracy) and most importantly because it follows the assumptions of the test. The variable is continuous and its values fit a normal distribution (Sig=.437) (see appendix 6).

Prosody was divided in four different categories when coded: a) Expression and Volume, b) Phrasing, c) Smoothness and d) Pace. Once agreed on the reliability of the data, a test was chosen to analyze prosody. In this case since the dataset is ordinal and the sample is not independent it was decided to use the Wilcoxon Signed-Rank test.

The assumptions for this test are: (1) The data is paired and come from the same population. (2) Each pair was randomly chosen. (3) The data is measured on an interval scale though it suffices that it is on an ordinal scale. All the assumptions are met by our data.

4.4.3 Quantitative analysis for the RQ2

The aim of the second research question is to answer whether the participants' emotional valence was positive (happiness, interest, excitement, etc.) or negative (sadness, anger, fear, etc.) during reading. To be able to do so, the Likert-scale questionnaires that were given to the students after reading each text were used. Using SPSS, the answers to the three questions that conformed the questionnaires were analyzed and compared. Since the dataset is ordinal and the sample is dependent, it was decided to use the Wilcoxon Signed-Rank test.

The questions are thought to give information about how the emotional valence was highlighted in the participants. The first question asks about the difficulty of the texts in a subjective way. The second question asks about their emotional valence (positive or negative) during the reading exercise and last, the third question asks about how their reading ability was perceived.

5 Results

5.1 RQ1: Is there a difference in fluency between reading in VR and on a piece of paper?

The results for this research question will be subdivided in the categories of fluency: accuracy, rate and prosody.

5.1.1 Accuracy

The scatter plot graph, in figure 3, shows the errors that the participants made while reading, without considering the categorization of the errors previously made. Based on the graph, the variables have a positive correlation, which means that the participants who performed well on paper also seem to perform well on VR. Besides, no outliers appear to be present in the dataset.

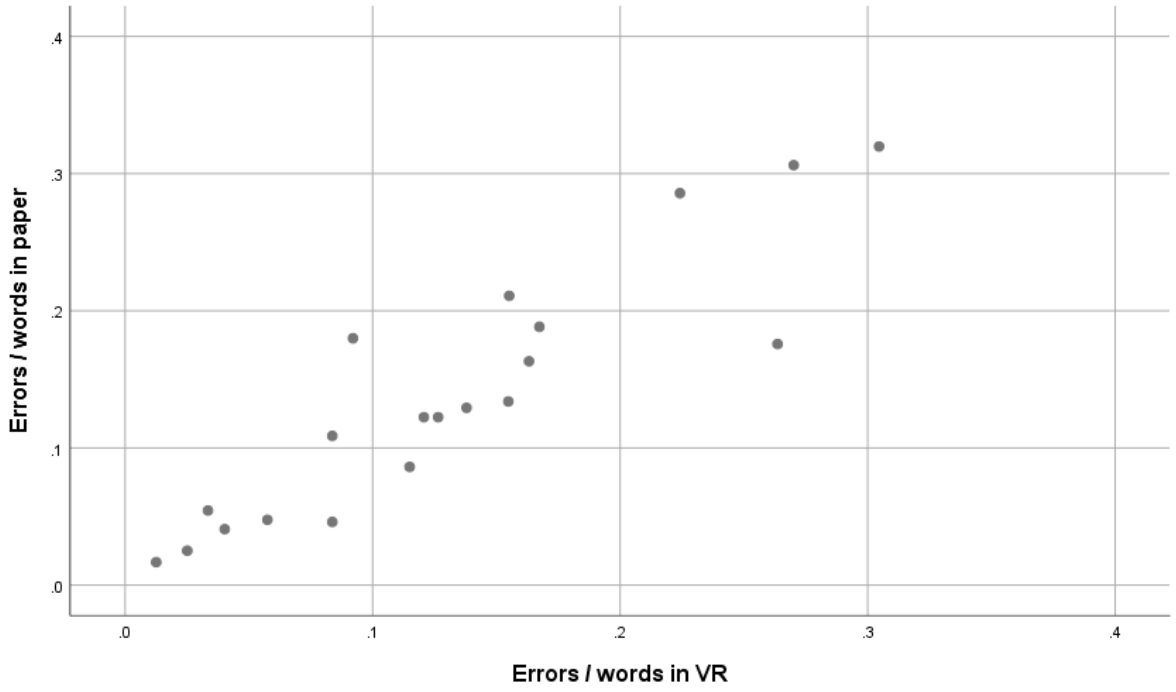


Figure 4: Scatter plot for Accuracy

As the data fits the assumptions for the test, a Paired-Sample T-Test was performed (see appendix 7), the p-value of the test is .446, which means there is not enough evidence to say there is a difference in the total errors between paper and VR. Because of that the errors were

looked in more detailed by analyzing each subcategory (substitution, insertion and omission) and running the tests independently.

Since the dataset has the same characteristics, it was decided to use a Paired-Sample T-Test, as long as our data keeps fitting the test's assumptions.

The Paired-Sample T-Test results, suggest that there is no statistically significant difference in the means of error rate for insertion, since the p-value is greater than 0.05. The same can be said for the results for the other error types, substitution and omission, which don't show statistical evidence that the means are different since the p-values are greater than 0.05, (see appendix 8).

5.1.2 Rate

Based on the Scatter Plot in figure 4, we can observe a positive correlation between the rate in paper and the rate in VR, which means the participants that performed well in paper also tend to perform well in VR. Besides, it can be observed that there are no outliers in the dataset.

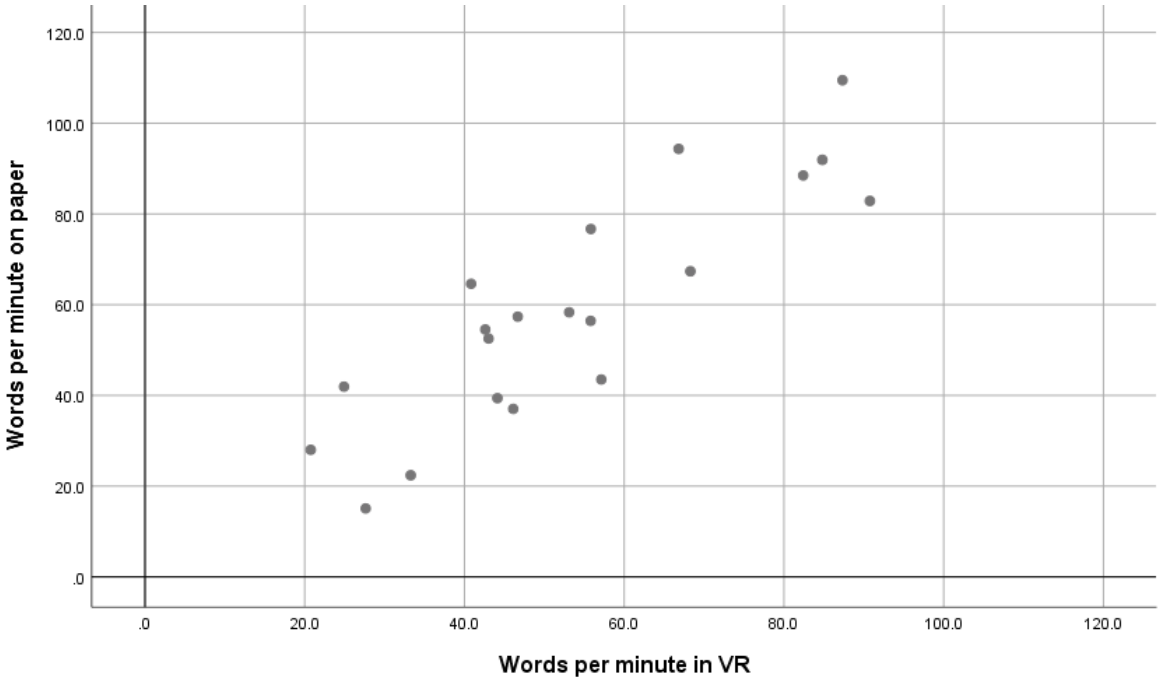


Figure 5: Scatter plot for Rate

The results of the Paired T-test (see appendix 9), suggest that there is not a statistically significant difference in the means of rate, since the p-value is greater than 0.05.

5.1.3 Prosody

According to the prosody stacked bar graphs in figures 6, 7, 8 and 9, the dataset seems to fit a normal distributioon. In the Expression and Volume, Smoothness and Pace graphs there can be observed that most of the participants performed well in both, VR and paper but with a slightly higher performance in VR, while in Phrasing it was evenly distributed.

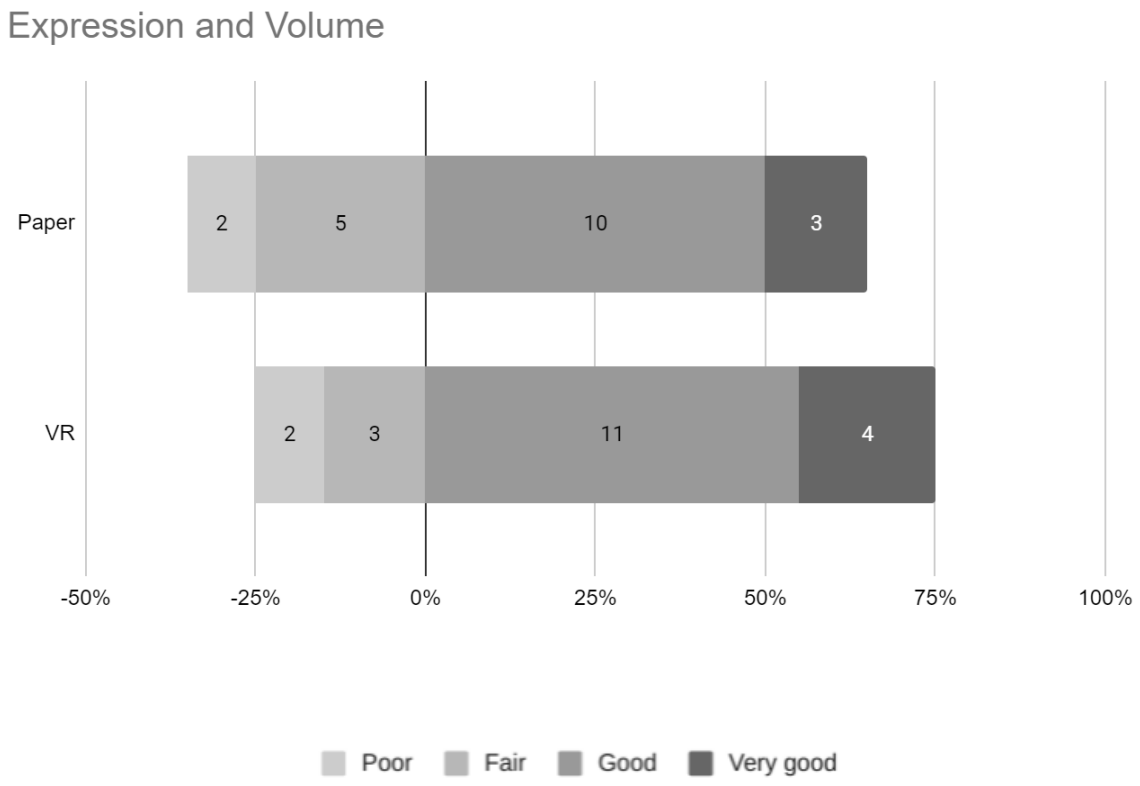


Figure 6: Expression and Volume Stacked bar

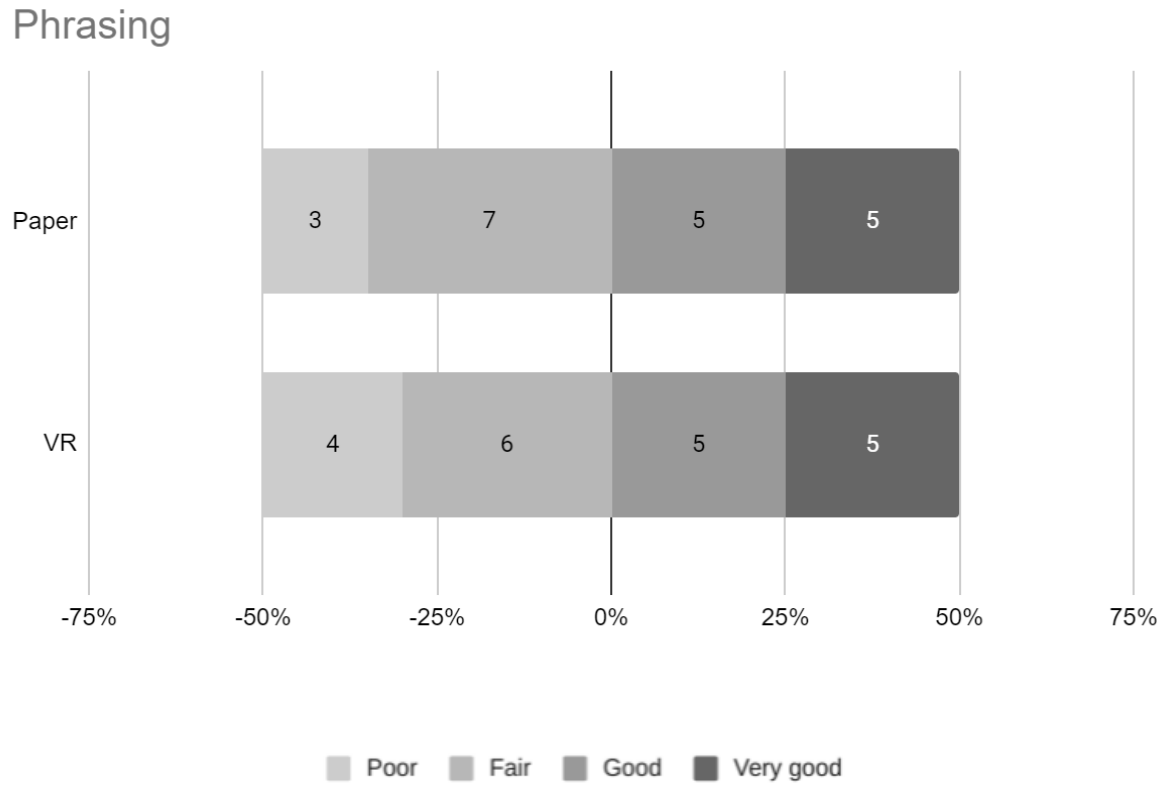


Figure 7: Phrasing Stacked bar

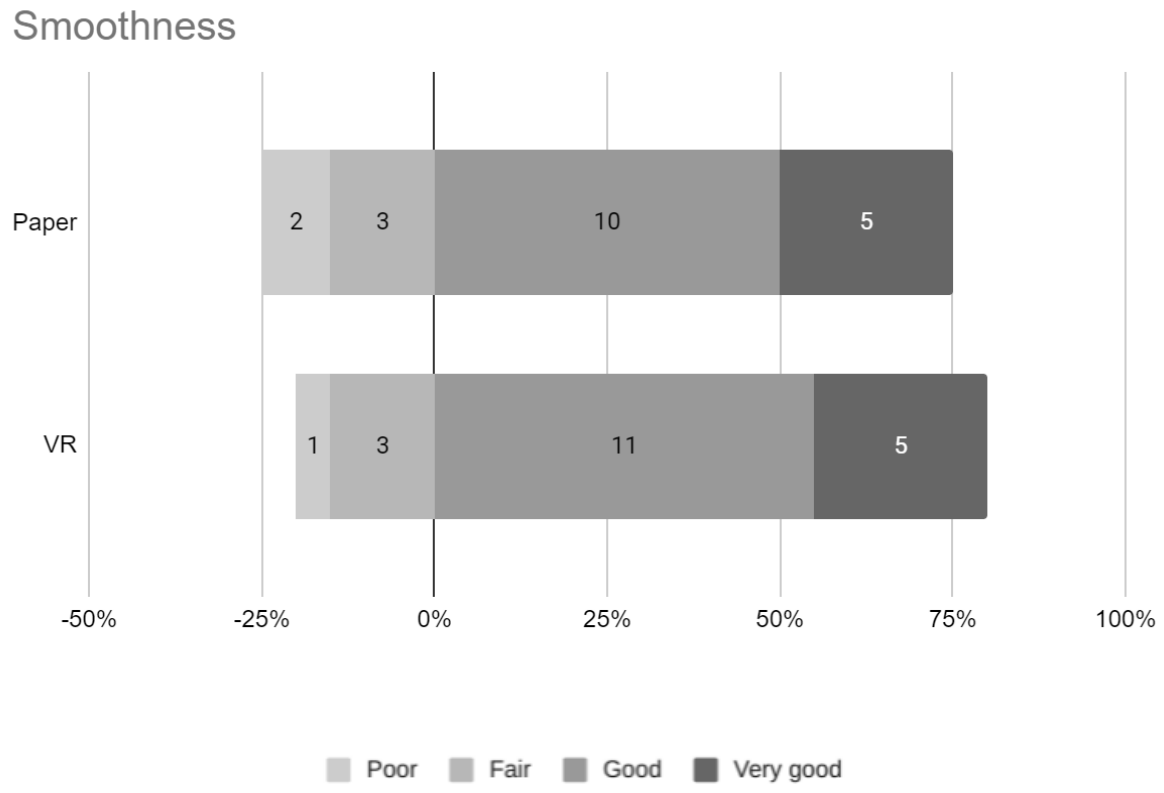


Figure 8: Smoothness Stacked bar

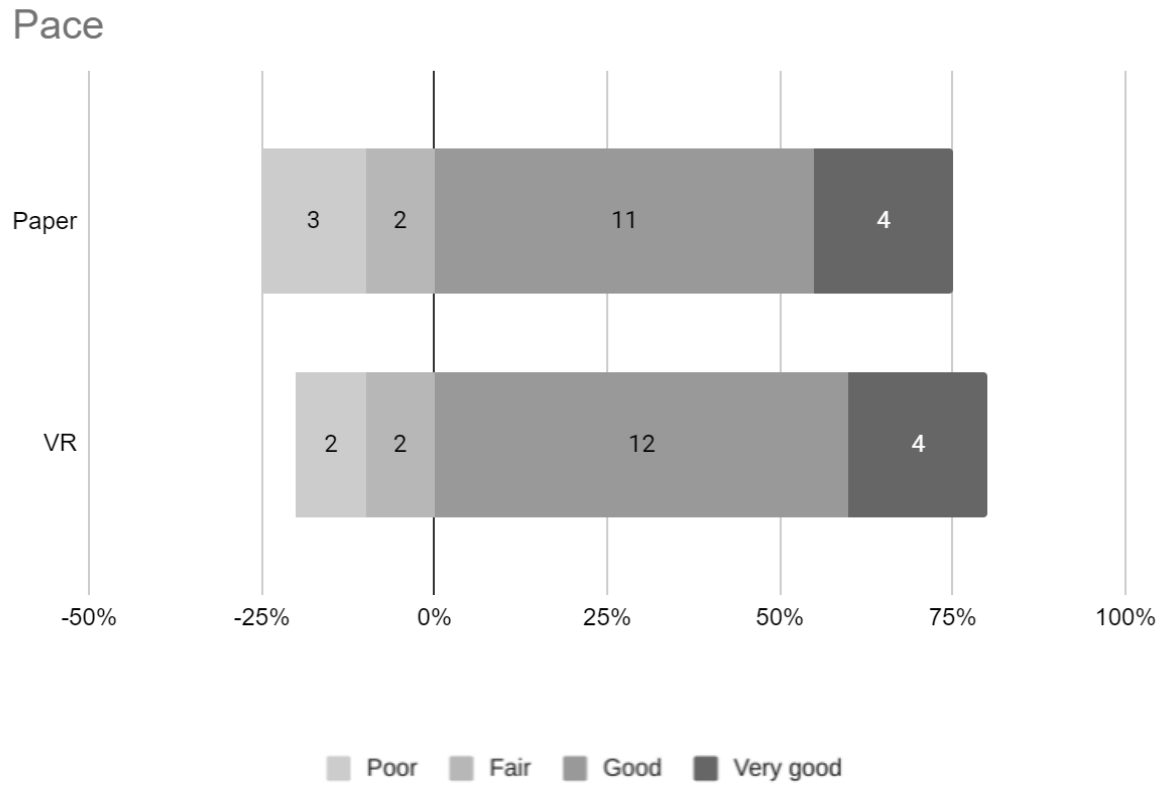


Figure 9: Pace Stacked bar

Even though it looked like the participants showed a slight improvement in VR in the graphs of Expression and Volume, Smoothness and Pace, the results from the Wilcoxon Signed-Rank test (see appendix 10) show that reading in VR instead of on paper did not elicit a statistically significant change in the students' performance in terms of prosody, (expression and volume: $Z = -1.732$, $p = .083$; phrasing: $Z = -.378$, $p = .705$; smoothness: $Z = -1.000$, $p = .317$; pace: $Z = -1.414$, $p = .157$).

5.2 RQ2: How positive and negative emotions were emphasized while reading?

As it can be observed in figure 10, it seems that the participants felt that VR was easier than paper, also in figures 11 and 12, it can be discerned that participants felt more positive emotions towards VR.

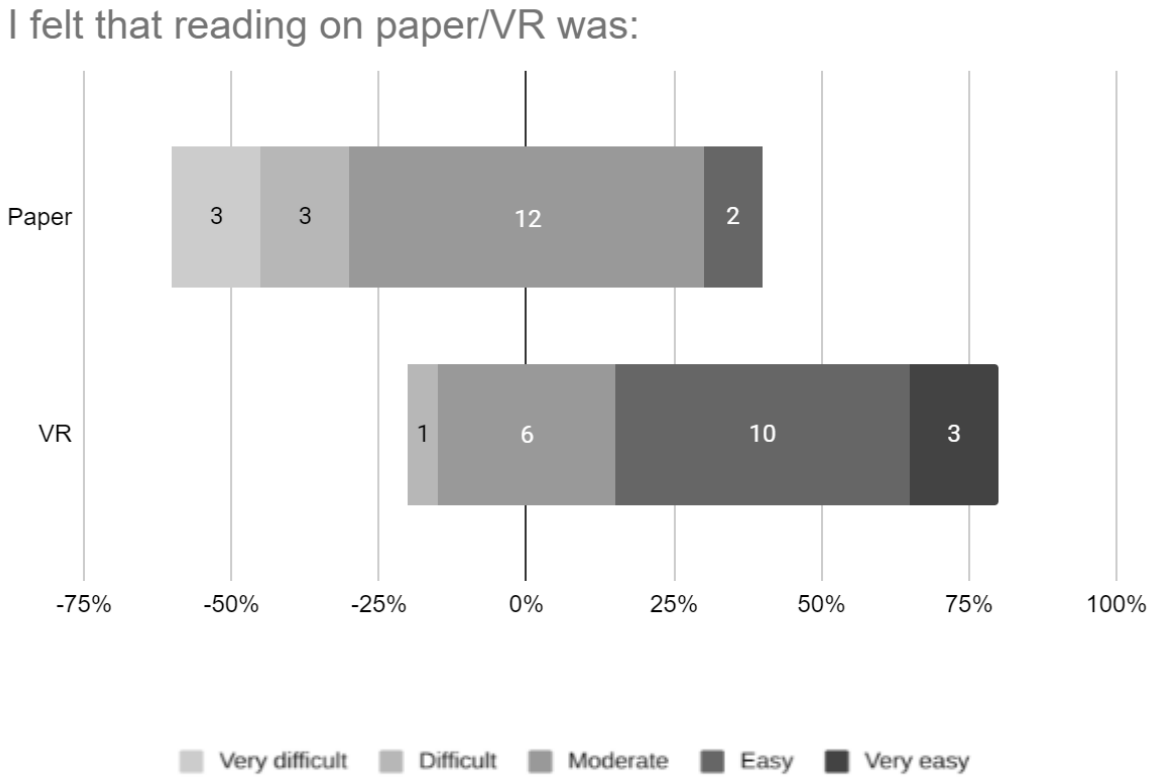


Figure 10: First question. Stacked bar graph

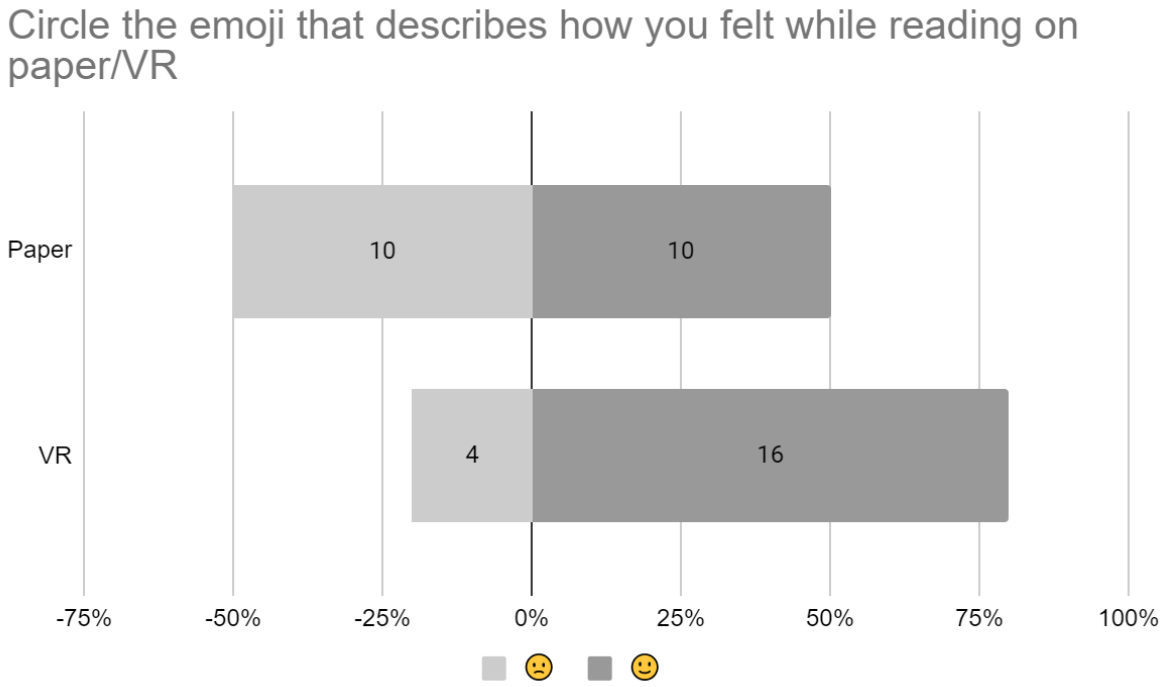


Figure 11: Second question. Stacked bar graph

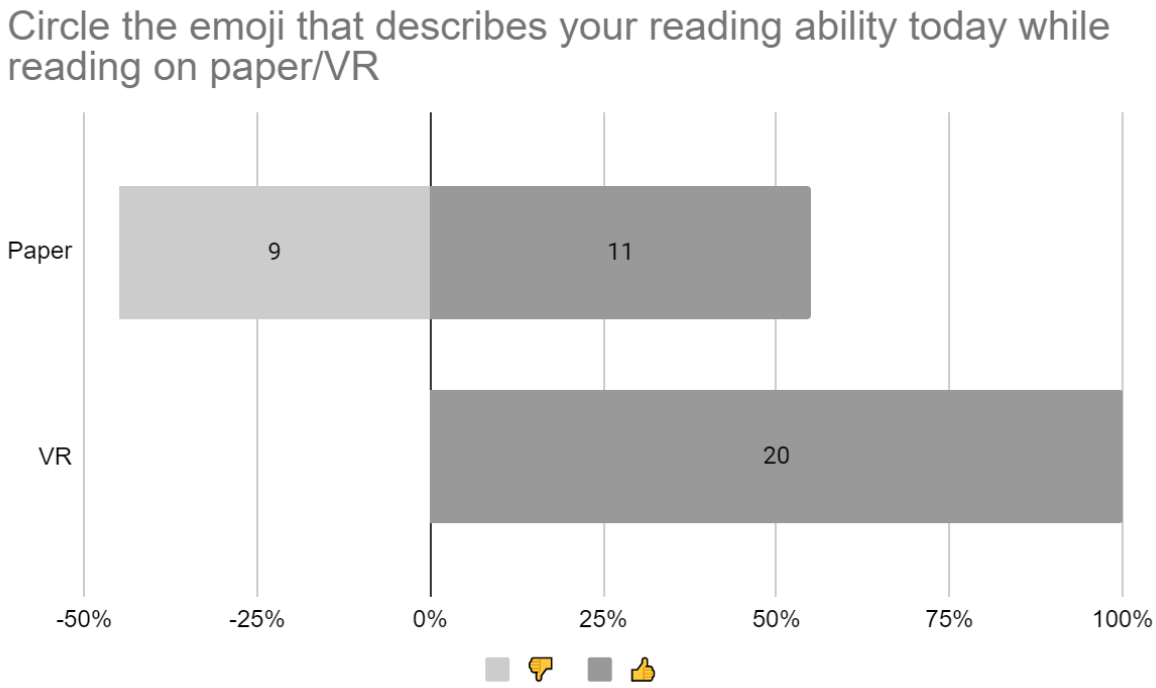


Figure 12: Third question. Stacked bar graph

According to the Wilcoxon Signed-Rank test results in table 3, there is a statistically significant difference between the students' emotions reading on paper and in the VR environment, since for the first, second and third question, Z=-4.119, p<.001; Z=-3.000, p=.003 and Z=-2.449, p=.014, respectively.

Table 6: Wilcoxon signed-rank test of Emotions

Test Statistics[a]

	Reading on VR was... - Reading on paper was...	Feeling while reading on VR - Feeling while reading on paper	Reading ability was... - Reading ability was...
Z	-4.119[b]	-3.000[b]	-2.449[b]
Asymp. Sig. (2-tailed)	.000	.003	.014
Exact Sig. (2-tailed)	.000	.004	.031
Exact Sig. (1-tailed)	.000	.002	.016
Point Probability	.000	.002	.016

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

5.3 Summary of the results

Summarizing the results, the first research question **(is there a difference in fluency between reading in VR and on a piece of paper?)**, the results of accuracy, rate and prosody did not elicit a statistically significant change in the students' performance.

As for the second research question (**How positive and negative emotions were empathized while reading?**) the results suggested that there is a statistically significant difference between students' emotions reading on paper and in VR.

6 Discussion

The general aim of this study is to research whether VR affects dyslexic students' fluency reading and also how it affects their emotions in terms of its valence. In this section, the results will be discussed.

Previous studies have researched the use of VR (Kalyvioti & Mikropoulos, 2014; Habib, et.al., 2012; Kalyvioti & Mikropoulos, 2012; etc.) and other technological affordances and ICT (Information and Communication Technologies) such as, digital pen technology (Chen, Tan & Lo, 2016), E-readers (Schneps, et.al., 2013) or mobile applications (Skiada, et.al., 2014) and its possible benefits for dyslexic students. But no research has been found that connects Virtual Reality, fluency and dyslexia. This might seem discouraging, but it is also an opportunity to find out if virtual environments ease up the reading challenge for dyslexic students and help them enjoy and read more fluently.

Something to point out, is that this research is a case study from an EdTech company, and since the sample is limited (N=23), more research will be necessary to delve into the topic and come up with some thorough results and conclusions. Moreover, this research is not trying to find a cure for dyslexia, especially where there is so much disagreement about the causes of dyslexia, its development and intervention. The aim, as explained, is to find out the preliminary results of the effects of VR in dyslexic students' fluency reading and also how it affects their emotional valence.

To answer the first research question **(is there a difference in fluency between reading in VR and on a piece of paper?)**, each element was analyzed individually. Accuracy, which also is comprised of three elements (insertion, substitution and omission), was analyzed as a whole and also individually and the results showed that there was no statistical evidence. Rate and prosody were also analyzed, and the results did not elicit a statistical difference between its means. Despite the lack of statistically significant results, theory suggests that VR has a positive effect on dyslexic children, for instance, in a research about visuospatial strengths of children with dyslexia using virtual reality and carried out by Attree, et.al. (2009), the results suggested that teenagers with dyslexia exhibit superior visuospatial strengths. These results could be very useful, because VR could be used then as an intervention method.

Moreover, other studies such as the one from Kalyvioti and Mikropoulos (2014) about virtual realities and dyslexia, affirmed that the relationship between virtual environments and dyslexia has been described several times in literature reviews and it can help not only dyslexic children per se but also it can serve as an orientation for family members, teachers and other professionals. It was also highlighted, in a paper from Leij (2013) about dyslexia and early intervention, that the use of information and new technologies, such as VR, can be useful in early intervention of students with dyslexia. Therefore, it could be agreed that further study with a wider data sample needs to be carried out for a future re-assessment of if and how VR as an impact on the fluency of students with dyslexia.

In addition, it is worth mentioning a study carried out by Schneps, et.al. (2013), that affirms that e-readers is a technological tool that facilitates reading and comprehension for students with dyslexia when compared to paper. This study is very interesting since the aim of this master's thesis is to compare another technological solution, such as VR, with paper, and although the study is not the same nor the technological solution, there are some similarities between them. For example, in the discussion of Schneps' study, it was mentioned that they believe the reason why e-readers facilitate better reading to dyslexic children is because of the possibility of enlarging or changing the font, as well as having fewer words per line than compared to a standard textbook, which is exactly the same as it can be done with VR, with the difference that VR has even more affordances that might have a positive effect on children's reading, such as immersiveness or sound.

More studies regarding the use of ICTs applied to dyslexia have mentioned and suggested the variation of the font type and size, colors on the screen, interactive multimedia instruments, etc. It should be emphasized the studied done by Zikl, et.al. (2015), which described the influence of fonts choice for reading difficulties. Their results revealed that using proper fonts show benefits in terms of accuracy and rate, two characteristics of fluency used also as measurement in this research, when it comes to students with dyslexia. Hence, a font adequate and specific for dyslexia was created. Therefore, even though for the sake of replicating what happened during Lyfta's workshop, special fonts were not used, this information and researchers are very important if we are to replicate either this master's thesis or one of the aforementioned studies.

However, not all the results are positive as not every person is the same, we cannot assume that the findings from the previous studies would affect everyone in the same way. A study done in

2012 by Habib, et.al., shows that some students with dyslexia preferred to read on paper rather than in VR. In the interviews the researchers carried out, the students pointed out that on paper it is easier to highlight or mark some explanations, and such. This is also quite interesting since brings a valuable point to take into account when it comes to the viability of introducing VR into classrooms and teaching. As it was mentioned, not every person is the same and not every student learns the same way. Moreover, in my opinion, the use of technology in teaching should always come with a purpose, should be tied to a curriculum with its learning objectives and should always bring some value to the learning.

The results from the second research question (**How positive and negative emotions were empathized while reading?**) suggested that there is a statistically significant difference between students' emotions reading on paper and in VR.

This master's thesis hypothesis was that VR would make a positive difference in terms of fluency on dyslexic children's reading, but after analyzing all the data and having studied the results, my hypothesis changed since there were no statistical difference on the results of the participants reading in VR compared to paper. Therefore, it is hypothesized, based on the results, that having positive emotions is enough to help students suffering from dyslexia boost their motivation and by doing so, increase their reading time and practice and love for reading, possibly helping to improve their reading skills and academic achievement.

Research has shown that students' emotions and academic achievement are connected (Goetz & Hall, 2013; Pekrun & Linnenbrink-García, 2014; Zeidner, 1998). Citing Pekrun (2002, p.149), "Positive emotions are essential for human behavior and adaptation. They help envision goals and challenges, open the mind to thoughts and problem-solving, protect health by fostering resiliency (...)". However, most educational psychology and educational research in general about emotions' effects on learning, students and teachers have been focused primarily on specific negative emotions, neglecting somehow positive emotions (Pekrun, et al., 2002).

Nonetheless, in the last decades, more and more studies have started focusing on positive emotions and their impact on academic achievement, self-regulation or motivation (Glaser-Zikuda, et al., 2005; Goetz, et.al., 2006; Pekrun, et al., 2009; Pekrun et al., 2010). They are thought to "preserve cognitive resources and focus attention on the learning task, support interest and intrinsic motivation, and facilitate deep learning (Pekrun, 2017, p.7)". The same

researcher discusses about the effect of opposite emotions (negative), such as boredom or sadness. He argues that they "reduce cognitive resources and task-related attention, to undermine both intrinsic and extrinsic motivation, and to promote shallow information processing (Pekrun, 2017, p.7).

Studies carried out to study students' emotional experiences result into two areas, students' self-regulation and motivation and students' learning and achievement (Schutz, et. al, 2010). It has been observed that a student with anxiety and fear about an activity may look for a teacher's help, looking for external regulation. (Linnenbrink, 2007; Pekrun, et.al., 2004). Studies (Linnenbrink, 2007; Meyer and Turner, 2007; Pekrun, et.al., 2002) also show that pleasant emotional experiences are usually associated with self-regulation, intrinsic motivation, and a tendency to engage meta-cognitive and/or self-regulatory learning strategies.

The above-mentioned author, Pekrun, has been one of the researchers that have focused on positive emotions the most, specifically on emotions and academic achievement. He proposes in his *Control Value Theory* that achievement emotions ("emotions related to achievement activities and their success and failure outcomes" (Pekrun, et al., 2017, p.6)), have an impact on students' cognitive resources and motivation, which at the same time are impacting their academic achievement. (Pekrun, & Linnenbrik-García, 2012). Particularly, positive activating emotions such as enjoyment or happiness, are thought to support interest, trigger their intrinsic and extrinsic motivation and facilitate deep learning.

Positive emotions, no matter their activation dimension status, besides facilitating academic achievement, are related to a series of motivational related variables, being competency beliefs, task and achievement value (Frenzel et al., 2007), intrinsic motivation (Ouano, 2011), self-regulated learning (Titz, 2001, cited in Pekrun, et al., 2002; Järvelä, 2011) or self-efficacy (Bandura, 1997; Schunk, 1991). In addition, whereas positive emotions have a great impact on academic achievement, they are also relevant on health, subjective well-being and lifelong learning (Pekrun, 2006).

Therefore, no matter the results for the first research question, it is clear that having a positive emotional valence is essential and has a profound impact on students' self-regulation and motivation and academic achievement. And this is important for every student, but it is especially meaningful for students with dyslexia, since as it was discussed in the literature

review, emotional well-being is essential for dyslexic students (Reid, 2011). Moreover, the same author points out the difficulty to recognize their own academic and non-academic success (Reid, 2011), which is related to self-efficacy, discussed previously. Furthermore, motivation is a keyword, when it comes to the well-being and academic achievement of students with dyslexia, since it decreases diminishing their emotional development and increasing their stress levels and anxiety (Mueller, 1976). Hence, the importance of keep digging in and finding more results that support this theories and lead to a better knowledge of how positive emotions affect academic achievement and how it can be beneficial for students with dyslexia.

7 Evaluation

Validity, reliability and the ethicality of a research have always been major concerns for researchers in both, quantitative and qualitative studies (Golafshani, 2003). Following Golafshani's line, Yin (2009) affirms that "the quality of a research study is judged in terms of its validity and reliability" (Yin, 2009, p. 32-33).

In this section the ethical issues, reliability and validity of the thesis will be assessed and critically analyzed.

7.1 Ethical issues

The University of Oulu's guidelines were consulted (RCR guidelines), as well as the ones from the Finnish Advisory Board on Research Integrity (TENK) to ensure that the research followed proper ethical directions.

Before conducting the data collection, a written consent was sent to the minors participating so it could be signed and agreed by their guardians, with explanatory text of how the data collection would work, in "common" language. Moreover, it was explained to them how the data collection would be used. The participation was completely voluntary, and they could regret at any moment if they didn't feel comfortable with the study. Moreover, the participants were not alone with the researchers during the data collection, one of their teachers were with them at all times. The minors were not exposed to any harm and were treated with respect and care.

All the information was inserted into Google Sheets and NVivo software, which are only for the researcher's use and for the present study. The data has been anonymized and it is available and enough for a future replication of the study. Since the video recordings cannot be anonymized and the data in those videos has already been quantified, will be deleted for the protection of the participants. No names were disclosed in this research.

7.2 Validity and Reliability

Transparency is a main priority in this study and has been achieved through a detailed and clear description of the research process. During the data collection, a strict script was followed in order to explain the task to all the participants in the same way and avoiding then possible bias.

Moreover, the questionnaires were read at loud for an easier understanding, considering that all the participants had a certain level of dyslexia.

Furthermore, it is worth mentioning researchers have an influence in the sample and population studied (Yow, 2006), this could be translated in the way the information is read and explained to the participants for example.

There are three main types of validity, according to Heale and Twycross (2015), which are content validity, construct validity and criterion validity, as shown in table 5 (Healy & Twycross, 2015, p.66).

Table 7: *Types of validity*

Type of validity	Description
Content validity	The extent to which a research instrument accurately measures all aspects of a construct
Construct validity	The extent to which a research instrument (or tool) measures the intended construct
Criterion validity	The extent to which a research instrument is related to other instruments that measure the same variables

All three types of validity have been taken into consideration when collecting and analyzing the data. Regarding the two first types of validity, content and construct validity, both, fluency and emotional valence were measured using video recording and questionnaires. These instruments were appropriate and accurate since they have been validated by conceptual frameworks and other professionals. Moreover, the instruments were able to measure in detail the constructs, in this case the participants emotional valence and participants fluency, which later were quantified and coded following proper methods and rubrics.

Addressing criterion validity, it is necessary to evaluate the types of instruments used and why they were chosen among others. To quantify accuracy, rate and evaluate prosody, a video recording method was chosen instead of quantifying the errors in situ, which would have decreased the stress produced by the camera in the participants, however it would have decreased extensively the accuracy of the error quantification and furthermore, it wouldn't have been possible to carry out a interrater reliability test, which would have made this study less reliable and trustworthy. For the second research question, to quantify emotions, a

questionnaire method was chosen instead of an interview of the participants to reduce their anxiety and keep it simple for the participants.

For the sake of this research, as it was mentioned in the data analysis section, and to ensure the reliability of the tests, an interrater test was done. The reason for the interrater test was the subjectivity of one of the analysis. The test carried out was a Cohens-Kappa, which gave a result of k=.81, that according to Fleiss guidelines (1981), shows an excellent agreement.

7.3 Limitations

Overall, there are certain limitations that should be acknowledged and considered for future research. Starting with the limited number of participants. As stated before, the lack of data sample might be the reason why there is not enough statistical difference to answer the first research question (is there a difference in fluency between reading in VR and on a piece of paper?).

Another limitation is the use of Likert-Scales for assessing students' academic emotions, since scales might not be the most trustworthy way of measuring emotions. There have also been some concerns about what the concept of academic emotions means across cultures (Bernando, Ouano & Salanga, 2009). Despite the concerns, the use of Likert-Scales as a measure of emotions in this master's thesis is a premeditated decision that has been reliable for this small-scale study. However, if a larger version of this study occurs, it is recommended a re-assessment of all the data collection and analysis methods for a better suit for a larger-scale research.

Last limitation could be the awareness of the learners while participating in the data collection. During the workshop Lyfta carried out, the intention was not to measure how the subject read and collect data, although there was a camera recording, the subject wasn't aware that his reading ability was being tested Therefore, during the data collection procedure for this master's thesis, the participants were well aware that the activity had to do with their reading ability and although they were told the reading activity was not an exam or was not going to make an effect of their school process, they fact that it was related to their reading ability had an effect of them and is a limitation of this study.

8 Conclusion

New technologies offer us new opportunities and different lenses which we can look through to a problem. It has been many years that researchers have been studying the whats, whys and hows of dyslexia. However, the new and fast-developing technology can help us attenuate the effects of dyslexia and give school children the same opportunities as their peers to grow and thrive in life.

With the present study, new technologies, such as VR, have been analyzed to study how the fluency of dyslexic students would be affected if they read in an virtual environment compared to a piece of paper, similar to the textbooks they usually have at school. This study has tried to simulate a case with specific characteristics (font or font size), but as Schneps, Thomson, Sonnert and Pomplun (2013) suggest in their study about the affordances of e-readers for dyslexic students, VR has many more affordances (fewer words per line, dynamic fonts, bigger font size, etc.) that could help students with dyslexia or other kinds of learning difficulties or disabilities, such as visual limitations.

Moreover, emotions have also played an important role in this study, as they play an essential role in students' everyday life. As Pekrun and Linnenbrink-García observe, "The classroom is an emotional place (…) Emotions are both *experienced* in the educational setting as well as *instrumental* for academic and personal growth" (2014, p.1). With this study, the aim has been to understand how the emotional valence was aroused on the participants while reading in the VR environment as well as on the piece of paper. Interestingly, there were positive significant results towards reading in the virtual environment since most of the students had more positive emotions while reading in VR.

These results have beneficial implications, especially for students with dyslexia, considering that positive emotions can facilitate the use of different learning strategies as well as promote self-regulation skills. Furthermore, they are thought to have an effect on students' intrinsic motivation (Mega, Ronconi & De Beni, 2013) and therefore, increase their learning achievement.

8.1 Future research

There are some recommended directions for future research based on this master's thesis, starting with a wider and more diverse sample size. As mentioned throughout this research, the sample size has not been large enough. This is important to take into consideration when designing a future research, not only to assure a fair representation of the population but also to limit the influence of outliers and assuring a significant difference (Patel, et.al., 2011). Moreover, it could be worth testing the same research with students whose mother tongue is not English, since English is a language with a deep and asymmetrical orthography (Sprenger-Charolles, et.al., 2013). Languages such as Spanish or Italian, are more transparent, with a more phoneme-grapheme correspondence than English (Pérez, 2005) and might be worth taking the study to this kind of languages. Therefore, a hypothesis is that, with a larger sample pool and a transparent language, there is a positive statistical difference towards the use of VR as a way to help dyslexic students improve their reading fluency. Also, the font used on the texts (both in VR and on paper) that the participants had to read during the data analysis, was a common font widely used in textbooks around the UK. It would be very interesting, as a future research, to use a font created specifically for dyslexic readers and use more VR feature, not only slides.

Lastly, regarding the second research question (how positive and negative emotions were empathized while reading?), an idea as a future research is to test specific emotions and other emotional dimensions such as approach and avoidance emotions, instead of just the emotional valence. This could give more knowledge about how specific emotions affect academic achievement, motivation and self-regulation on students with dyslexia. An example of specific emotional dimensions could be the ones that propose Pekrun and Linnenbrink-García (2012) on their *Control Value Theory*: positive activating (e.g., enjoyment, hope, pride), positive deactivating (e.g., relaxation, relief), negative activating (e.g., anger, anxiety, shame) and/or negative deactivating (e.g., boredom, hopelessness).

8.2 Practical implications

Continuing with this kind of research could make a huge impact in the lives of students with dyslexia. Even if no new intervention method is created, their motivation would still increase and their wellbeing improved since as it was described in the literature review section, dyslexic students often suffer from anxiety (Reid, 2011), and a positive emotional valence plays an

important role in students with dyslexia, since it is proven how they reduce depression and anxiety often reported in dyslexic individuals (Nalavani, et al., 2017).

Other applications of this kind of research might be in the world of education technology (EdTech) companies, since as companies as dyslexiefont.com have done, it would provide more information to those companies which are interested in creating a product for this kind of learning difficulty. Not to mention parents and teachers of dyslexic students. Research focused on dyslexia (or other learning difficulties) interventions, gives information, resources, hope and new methods to help their children and students improve their learning achievement.

References

Abedi, A., Sepanta, M., Yarmohammadian, A., Ghamarani, A., & Faramarzi, S. (2019). The Effect of Positive Emotion Training Program on Psychological Well-being of Students with Dyslexia. *Avicenna Journal of Neuro Psycho Physiology*, 5(3), 0-0.

Asan, O., & Montague, E. (2014). Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Informatics in primary care*, 21(4), 161.

Allington, R. L. (1983). The reading instruction provided readers of differing reading abilities. *The Elementary School Journal*, *83*(5), 548-559.

Attree, E. A., Turner, M. J., & Cowell, N. (2009). A virtual reality test identifies the visuospatial strengths of adolescents with dyslexia. *CyberPsychology & Behavior*, *12*(2), 163-168.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *The Journal of the Learning Sciences*, *17*(1), 102-141.

Bandura, A. (2010). Self-efficacy. *The Corsini encyclopedia of psychology*, 1-3.

Beck, M., Burns, M. K., & Lau, M. (2009). The effect of preteaching reading skills on the on-task behavior of children identified with behavioral disorders. *Behavioral Disorders*, *34*(2), 91-99.

Benfatto, M. N., Seimyr, G. Ö., Ygge, J., Pansell, T., Rydberg, A., & Jacobson, C. (2016). Screening for dyslexia using eye tracking during reading. *PloS one*, 11(12).

British Dyslexia Association. (2018). *About dyslexia*. Retrieved from https://www.bdadyslexia.org.uk/dyslexia/about-dyslexia

Boder, E. (1970). Developmental dyslexia: A new diagnostic approach based on the identification of three subtypes. *Journal of School Health*, *40*(6), 289-290.

Burden, R. (2008). Is dyslexia necessarily associated with negative feelings of self-worth? A review and implications for future research. *Dyslexia*, 14(3), 188-196.

Burns, M. K. (2007). Reading at the instructional level with children identified as learning disabled: Potential implications for response-to-intervention. *School Psychology Quarterly*, *22*(3), 297.

Burns, M. K., Hodgson, J., Parker, D. C., & Fremont, K. (2011). Comparison of the effectiveness and efficiency of text previewing and preteaching keywords as small-group reading comprehension strategies with middle-school students. *Literacy Research and Instruction*, *50*(3), 241-252.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., & Seidlová-Málková, G. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23(6), 678-686.

Carrozzino, M., & Bergamasco, M. (2010). Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage*, *11*(4), 452-458.

Charlton, B. C. (2005). *Informal assessment strategies: Asking questions, observing students, and planning lessons that promote successful interaction with text*. Pembroke Publishers Limited.

Chen, C. M., Tan, C. C., & Lo, B. J. (2016). Facilitating English-language learners' oral reading fluency with digital pen technology. *Interactive Learning Environments*, 24(1), 96-118.

Cidrim, L., & Madeiro, F. (2017). Information and Communication Technology (ICT) applied to dyslexia: literature review. *Revista CEFAC*, 19(1), 99-108.

Cox, A. R. (1985). Alphabetic phonics: An organization and expansion of Orton-Gillingham. *Annals of Dyslexia*, 35(1), 187-198.

Creswell, J. W. (2003). Research Design: Qualitative. *Quantitative, and mixed methods*.

Das, J. P. (2009). *Reading difficulties and dyslexia: An interpretation for teachers*. Sage Publications.

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into practice*, 30(3), 165-175.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. Reading & Writing Quarterly: *Overcoming Learning Difficulties*, 14(2), 135-163.

Fane, J., MacDougall, C., Jovanovic, J., Redmond, G., & Gibbs, L. (2018). Exploring the use of emoji as a visual research method for eliciting young children's voices in childhood research. *Early Child Development and Care*, 188(3), 359-374.

Franklin Institute (2019). *History of Virtual Reality*. Retrieved from https://www.fi.edu/virtual-reality/history-of-virtual-reality.

Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. *In the International Scientific Conference eLearning and Software for Education* (Vol. 1, No. 133, pp. 10-1007).

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, 17(5), 478-493.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, 127, 252-266.

Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy?. *British journal of educational technology*, 46(2), 412-422.

Galaburda, A. M., Menard, M. T., & Rosen, G. D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 91(17), 8010-8013.

Gickling, E. E., & Thompson, V. P. (1985). A personal view of curriculum-based assessment. *Exceptional Children*, 52(3), 205-218.

Gläser-Zikuda, M., Fuß, S., Laukenmann, M., Metz, K., & Randler, C. (2005). Promoting students' emotions and achievement–Instructional design and evaluation of the ECOLE-approach. *Learning and Instruction*, 15(5), 481-495.

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology*, 76(2), 289-308.

Goldberg, R. J., Higgins, E. L., Raskind, M. H., & Herman, K. L. (2003). Predictors of success in individuals with learning disabilities: A qualitative analysis of a 20-year longitudinal study. *Learning Disabilities Research & Practice*, 18(4), 222-236.

Götz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. *International guide to student achievement*, 192-195.

Guttorm, T.K., Leppänen, P.H.T., Richardson, U., & Lyytinen, H. (2001). Event-related potentials and consonant differentation in newborns with familial risk for dyslexia. *Journal of Learning Disabilitie*s, 34(6), 534-544.

Habib, L., Berget, G., Sandnes, F. E., Sanderson, N., Kahn, P., Fagernes, S., & Olcay, A. (2012). Dyslexic students in higher education and virtual learning environments: an exploratory study. *Journal of Computer Assisted Learning*, 28(6), 574-584.

Hales, G (2009). Self.esteem and counselling. In G.Reid (Ed.), *The routledge Companion Book of Dyslexia*. London: Routledge.

Hall, L., Hume, C., & Tazzyman, S. (2016, June). Five degrees of happiness: Effective smiley face likert scales for evaluating with children. In *Proceedings of the 15th International Conference on Interaction Design and Children* (pp. 311-321). ACM.

Kalka, D., & Lockiewicz, M. (2018). Happiness, life satisfaction, resiliency and social support in students with dyslexia. *International Journal of Disability, Development and Education*, 65(5), 493-508.

Harrison, C. (2003*). Understanding reading development*. Sage.

Hettinger, L. J., & Riccio, G. E. (1992). Visually induced motion sickness in virtual environments. Presence: *Teleoperators & Virtual Environments*, 1(3), 306-310.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, 30(4), 461-473.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how?. *The Reading Teacher*, 58(8), 702-714.

Hulme, C. E., & Snowling, M. E. (1994). Reading development and dyslexia. *In The chapters in this book are based on papers presented at the 3rd International Conference of the British Dyslexia Assn entitled'Dyslexia: Towards a Wider Understanding," held in Manchester, England, Apr 1994*. Whurr Publishers.

International Dyslexia Association. (2017) *Dyslexia at a Glance*. Retrieved from https://dyslexiaida.org/

Jarvela, S. (Ed.). (2011). *Social and emotional aspects of learning*. Elsevier.

Johnson, M., & Peer, L. (Eds.). (2002). The Dyslexia Handbook, 2002: *A Compendium of Articles, Checklists, Resources and Contacts for Dyslexic People, Their Families and Teachers*. British Dyslexia Association.

Kaye, L. K., Malone, S. A., & Wall, H. J. (2017). Emojis: Insights, affordances, and possibilities for psychological science. *Trends in cognitive sciences*, *21*(2), 66-68.

Kaye, L. K., Wall, H. J., & Malone, S. A. (2016). "Turn that frown upside-down": A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior*, *60*, 463-467.

Kalyvioti, K., & Mikropoulos, T. A. (2012). Memory performance of dyslexic adults in virtual environments. *Procedia Computer Science*, *14*, 410-418.

Kalyvioti, K., & Mikropoulos, T. A. (2014). Virtual environments and dyslexia: A literature review. *Procedia Computer Science*, 27, 138-147.

Kennedy, R. S., Drexler, J., & Kennedy, R. C. (2010). Research in visually induced motion sickness. *Applied ergonomics*, 41(4), 494-503.

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: immersion aids recall. *Virtual Reality*, *23*(1), 1-15.

Lambert, C., & Kormos, J. (2014). Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language Acquisition. *Applied Linguistics*, *35*(5).

LaValle, S. (2017). Virtual reality. *Cambridge.*

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424-1442.

Leppänen, P.H.T., Pihko, E., Eklund, K.M.& Lyytinen, H. (1999) Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects, *Neuroreport* 10 (5), 967–973.

Leppänen, P.H.T., Richardson, U., Pihko, E., Eklund, K.M., Guttorm, T.K., Aro, M., & Lyytinen, H. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Developmental Neuropsychology*, 22(1), 407-422.

Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In *Emotion in education* (pp. 107-124). Academic Press.

Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulti*es, 23(2), 107-135.

Livingstone, M. S., Rosen, G. D., Drislane, F. W., & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences,* 88(18), 7943-7947.

Lyytinen, P. & Lyytinen, H. (2004). Growth and predictive relations of vocabulary and inflectional morphology in children with and without familial risk for dyslexia. *Applied Psycholinguist*ics, 25, 397–411.

Lyytinen, P., Eklund, K. & Lyytinen, H. (2005). Language development and literacy skills in late-talking toddlers with and without familial risk for dyslexia. *Annals of Dyslexia*, 55, 2, 166-192.

Marsh, J. (Ed.). (2005). *Popular culture, new media and digital literacy in early childhood*. Psychology Press.

Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual technologies trends in education. *EURASIA Journal of Mathematics Science and Technology Education*, 13(2), 469-486.

Mei, H. H., & Sheng, L. S. (2011). Applying situated learning in a virtual reality system to enhance learning motivation. *International journal of information and education technology*, 1(4), 298-302.

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of educational psychology*, 106(1), 121.

Meyer, D. K., & Turner, J. C. (2007). Scaffolding emotions in classrooms. In *Emotion in education* (pp. 243-258). Academic Press.

Miles, T. R. (Ed.). (2004). *Dyslexia and stress* (Vol. 21). Wiley.

Miller, J. F. (2018). *Do You Read Me?: Learning Difficulties, Dyslexia and the Denial of Meaning*. Routledge.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56(3), 769-780.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., ... & Tóth, D. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, 29, 65-77.

Mueller, J. H. (1976). Anxiety and cue utilization in human learning and memory.

Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental brain research,* 235(3), 889-901.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Nalavany, B. A., Carawan, L. W., & Brown, L. J. (2011). Considering the role of traditional and specialist schools: do school experiences impact the emotional well-being and self-esteem of adults with dyslexia?. *British Journal of Special Education*, 38(4), 191-200.

Nalavany, B. A., Logan, J. M., & Carawan, L. W. (2018). The relationship between emotional experience with dyslexia and work self-efficacy among adults with dyslexia. *Dyslexia*, 24(1), 17-32.

NHS. (2018). *Overview. Dyslexia*. Retrieved from https://www.nhs.uk/conditions/dyslexia/

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research?. *Cognition*, 35(2), 159-182.

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

OpenDyslexic (cited 2019 Oct 18). Available from: https://opendyslexic.org/about

Ouano, J. A. (2011). Motivational antecedents of academic emotions in Filipino college students. *The Asia-Pacific Education Researcher*, *20*(1), 127-132.

Palti.G. (2002). Social and Emotional Aspects of Dyslexia. *The Dyslexia Handbook 2002*, 191-195.

Parker, J. D., Creque Sr, R. E., Barnhart, D. L., Harris, J. I., Majeski, S. A., Wood, L. M., ... & Hogan, M. J. (2004). Academic achievement in high school: does emotional intelligence matter?. *Personality and individual differences*, 37(7), 1321-1330.

Parsons, T. D., Rizzo, A. A., Rogers, S., & York, P. (2009). Virtual reality in paediatric rehabilitation: a review. *Developmental neurorehabilitation*, 12(4), 224-238.

Passig, D. (2011). The impact of Immersive Virtual Reality on educator's awareness of the cognitive experiences of pupils with dyslexia. *Teachers College Record*, 113(1), 181-204.

Peachey, A. (2010). The third place in second life: Real life community in a virtual world. In *Researching learning in virtual worlds* (pp. 91-110). Springer, London.

Pedroli, E., Padula, P., Guala, A., Meardi, M. T., Riva, G., & Albani, G. (2017). A psychometric tool for a virtual reality rehabilitation approach for dyslexia. *Computational and mathematical methods in medicine*, 2017.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, *18*(4), 315-341.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of educational Psychology*, *101*(1), 115.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, *102*(3), 531.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, 37(2), 91-105.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Positive emotions in education.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child development*, *88*(5), 1653-1670.

Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and validation of the Test Emotions Questionnaire (TEQ). *Anxiety, Stress & Coping*, 17(3), 287-316.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (pp. 259-282). Springer, Boston, MA.

Pekrun, R., & Linnenbrink-Garcia, L. (2014). Introduction to emotions in education. In *International handbook of emotions in education* (pp. 11-20). Routledge.

Pellas, N., Kazanidis, I., Konstantinou, N., & Georgiou, G. (2017). Exploring the educational potential of three-dimensional multi-user virtual worlds for STEM education: A mixed-method systematic literature review. *Education and Information Technologies*, 22(5), 2235-2279.

Pink, S. (2003). Interdisciplinary agendas in visual research: re-situating visual anthropology. *Visual studies*, *18*(2), 179-192.

Thomson, R., & McLeod, J. (2015). New frontiers in qualitative longitudinal research: an agenda for research. *International Journal of Social Research Methodolo*gy, 18(3), 243-250.

Rasinski, T. (2004). Creating fluent readers. *Educational Leadership*, *61*(6), 46-51.

Reid, G. (2005). *Dyslexia: a complete guide for parents*. John Wiley & Sons.

Reid, G. (2011). *Dyslexia*. A&C Black.

Reynolds-Keefer, L., & Johnson, R. (2011). Is a picture is worth a thousand words? Creating effective questionnaires with pictures. *Practical Assessment, Research & Evaluation*, *16*.

Richardson, U., Leppänen, P.H.T., Leiwo, M., & Lyytinen, H. (2003). Speech perception of infants with high familial risk for dyslexia differ at the age of 6 months. *Developmental Neuropsycholo*gy, 3, 385-397

Riddick, B. (2006). Dyslexia friendly schools in the UK. *Topics in Language Disorders*, 26(2), 144-156.

Riddick, B., Farmer, M., Serling, C. (1997). *Students and Dyslexia: Growing up with a Specific Learning Difficulty*. London: Whurr.

Rizzo, A., John, B., Newman, B., Williams, J., Hartholt, A., Lethin, C., & Buckwalter, J. G. (2013). Virtual reality as a tool for delivering PTSD exposure therapy and stress resilience training. *Military Behavioral Health*, 1(1), 52-58.

Rosemblum, L. J., & Cross, R. A. (1997). The challenge of virtual reality. *Visualization & modeling*, 325-399.

Sakamat, N., Sabri, S. N., & Diah, N. M. (2017). Multimedia elements of digital storytelling for Dyslexic children. *Scientific Research Journal*, 14(2), 35-48.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child development*, 61(6), 1728-1743.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological psychiatry*, *57*(11), 1301-1309.

Sherman, W. R., & Craig, A. B. (2003). Understanding virtual reality. *San Francisco, CA: Morgan Kauffman*.

Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social science information*, *44*(4), 695-729.

Schneider, D. (2019). The Impact of Immersive Virtual Reality on Learning, Post-Hoc; a Cautionary Tale. *Issues and Trends in Educational Technology*, 7(1).

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*, *12*(3), 177-186.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into practice*, *30*(3), 158-164.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational psychologist*, *26*(3-4), 207-231.

Schutz, P. A., Quijada, P. D., De Vries, S., & Lynde, M. (2010). *Emotion in educational contexts.*

Schnall, S. (2010). Affect, mood and emotions. *Social and emotional aspects of learning*, 59-64.

Schneps, M. H., Thomson, J. M., Chen, C., Sonnert, G., & Pomplun, M. (2013). E-readers are more effective than paper for some with dyslexia. *PloS one*, 8(9), e75634.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annu. Rev. Psychol.*, 59, 451-475.

Shuman, V., & Scherer, K. R. (2014). Concepts and structures of emotions. In *International handbook of emotions in education* (pp. 23-45). Routledge.

Sigmundsson, H. (2005). Do visual processing deficits cause problem on response time task for dyslexics?. *Brain and cognition*, *58*(2), 213-216.

Southgate, E., Smith, S. P., Cividino, C., Saxby, S., Kilham, J., Eather, G., ... & Bergin, C. (2019). Embedding immersive virtual reality in classrooms: Ethical, organisational and

educational lessons in bridging research and practice. *International journal of child-computer interaction*, 19, 19-29.

Snowling, M. J., & Hulme, C. (2012). Annual Research Review: The nature and classification of reading disorders–a commentary on proposals for DSM-5. *Journal of Child Psychology and Psychiatry*, 53(5), 593-607.

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, 142(5), 498.

Snowling, M., & Thomson, M. (1992). Dyslexia—integrating theory and practice. *Child Language Teaching and Therapy*, 8(2), 212-214.

Sprenger-Charolles, L., Colé, P., & Serniclaes, W. (2013). *Reading acquisition and developmental dyslexia*. Psychology Press.

SPSS (cited 2019 May). Available from https://www.spss-tutorials.com/spss-paired-samples-t-test/

Stein, J. (2003). Visual motion sensitivity and reading. *Neuropsychologia*, 41(13), 1785-1793.

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in neurosciences*, 20(4), 147-152.

Skiada, R., Soroniati, E., Gardeli, A., & Zissis, D. (2014). EasyLexia: A mobile application for children with learning difficulties. *Procedia Computer Science*, 27, 218-228.

Szadokierski, I., Burns, M. K., & McComas, J. J. (2017). Predicting intervention effectiveness from reading accuracy and rate measures through the instructional hierarchy: Evidence for a skill-by-treatment interaction. *School Psychology Review*, *46*(2), 190-200.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and language*, 9(2), 182-198.

Tallal, P. (2000). The science of literacy: From the laboratory to the classroom. *Proceedings of the National Academy of Sciences*, 97(6), 2402-2404.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, 682(1), 27-47.

Turner, J. E., & Waugh, R. M. (2007). A dynamical systems perspective regarding students' learning processes: Shame reactions and emergent self-organizations. In *Emotion in education* (pp. 125-145). Academic Press.

Van der Leij, A. (2013). Dyslexia and early intervention: what did we learn from the Dutch Dyslexia Programme?. *Dyslexia*, 19(4), 241-255.

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

ViWo (2009). *1st Special Track on Virtual Worlds for academic, organizational, and life-long learning.*

Wu, F., Liu, Z., Wang, J., & Zhao, Y. (2015, March). Establishment virtual maintenance environment based on VIRTOOLS to effectively enhance the sense of immersion of teaching equipment. *In 2015 International Conference on Education Technology, Management and Humanities Science* (ETMHS 2015). Atlantis Press.

Yuasa, M., Saito, K., & Mukawa, N. (2011). Brain activity when reading sentences and emoticons: an fMRI study of verbal and nonverbal communication. *Electronics and Communications in Japan*, *94*(5), 17-24.

Zeidner, M. (2010). Test anxiety. *The Corsini encyclopedia of psychology*, 1-3.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., ... & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological science*, 21(4), 551-559.

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

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, *30*(3), 211-217.

Appendix 1

HOME WHAT IS IT? VIDEOS STORYWORLDS BLOG ABOUT AGENCY [FI] LOGIN

A REAL SURPRISE IN VIRTUAL REALITY.

14 December, 2017

Hi, I'm María Carrasco. I'm a qualified teacher working with a Finnish Education Technology company called Lyfta, part-time, while continuing my Master's degree in Learning, Education and Technology at the University of Oulu in Finland.

When introducing Lyfta to children, we're immersing them in story-worlds to capture their imaginations and help make it easier for teachers to delve into complex topics. We're always surprised by the results this can give, but we could never have imagined the impact it would have on one student in particular, when my colleague Serdar Forik visited a school in North West London, while running a Virtual Reality based workshop with a group of Year 8 children.

What happened?

There were 6 children in the group, using one VR headset between them. Serdar asked the children to describe what they could see and hear while wearing the headset with their classmates, to keep the whole group engaged throughout the session. Kit was the second child to have a go – during which he explored the weaving factory in a small Ethiopian village called Awra Amba.

Serdar asked Kit to click on some of the available multimedia content and read out the text for everyone to hear.

What Serdar didn't know, was that Kit is severely dyslexic. Kit hesitated for a moment, but then went on to read the short passage in front of his peers – which he delivered perfectly. When he took the headset off, he told Serdar that that was the first time he had read aloud without making a mistake.

Later that day, Serdar discussed what had happened with Kit's form tutor and interviewed Kit to find out more about his experience. Kit's mum, Annabel Cody (a journalist who had previously written a newspaper article about Kit's dyslexia) wrote to Serdar after Kit told her what had happened. Annabel and Serdar met a couple of times to discuss further, and Serdar also interviewed Annabel – now quite an expert in the field of dyslexia, having read around the subject for a number of years.

While our team at Lyfta has significant expertise in UX design – with particular effort placed on making text as clear and readable as possible – this was a totally unexpected outcome that caught everybody by surprise.

After hearing this story, I was inspired to explore the topic further. I transcribed and analysed Kit's interview and spoke to my supervisor at university. I had to make an application to change the focus of my Master's thesis, so that I could conduct a formal study about the effects of VR learning environments on reading for children with dyslexia. This was approved, and we have now designed a study with Serdar and my supervisor, which will be ready to conduct in 2018.

An academic study

We have decided to work with English text and would like to conduct the study with dyslexic students from the UK. We will compare similar texts in textbooks and VR environments to see if there is a significant difference in the way the children respond.

We are looking for 50 pupils:

- aged 10-14
- who are generally academically able
- for whom English is a first language
- who have dyslexia and find it difficult to read out loud.

We would like to work with 5 host schools around England and Wales, where we can be stationed for a day (in each school), with around 10 children visiting from other schools in the area/region.

Can you help us with the study?

1. We are looking for host schools around England and Wales (i.e. the South of England, London, the Midlands, the North and in South Wales). The schools should be easily accessible for pupils who may travel from around the region / from other schools. Please get in touch if you feel that your school can act as a hub in your area and you can lend us a classroom for a day.

2. We are looking for children with a similar profile to Kit, as stated above. If you have a pupil who you feel matches the profile, please discuss this idea with their parents and let us know if they're happy to be part of the study.

3. We would love support/advice to help us with the study. In order to be able to conduct the research in a timely manner, Lyfta has agreed to pay for the costs of the study (including equipment costs, my time, flights, travel and accommodation). As Lyfta is a start-up, with very limited funds, our budget is relatively small, so we would very much appreciate any support (e.g. from individuals/foundations/organisations who may be able to help/advise with regards to funding).

The experience with Kit was a complete surprise for all of us. We are very excited by the possibilities and look forward to working with primary and secondary schools to delve into the topic and see what we can find – which we hope can be built on to contribute to improving the lives of people with dyslexia in years to come.

Please fill in the short form below if you'd like to help us with the study:

Name *

School / organisation name *

Email Address *

I may be able to help, because: *

(please check one or more boxes)

- I know a pupil who may match the profile described
- Our school may be able to host part of the study
- I/we may be able to offer other support/advice
- Other

Message

SEND

Appendix 2

1972

Awra Amba was founded by 19 farmers who came together from different villages.

At the time, the community did not have any shared land. People were scattered around in different areas, all living off their own land.

1986

The families who had started the Awra Amba ideology, moved to the same village in order to work together and practice the principles with one another.

At the time there were a total of 81 people living on 40 hectares of land.

1988

The community had grown to 159 people, when problems started. Some people did not like our ideas and way of life and we were violently attacked during the night, forcing everyone to flee Awra Amba.

For six years, the community lived in exile in Southern Ethiopia.

1994

When we returned to Awra Amba six years later, there were 255 of us. We came back to find out that our land had been reallocated to other farmers by the Government during a land reform.

But a kind farmer gave us 1.5 hectares of his land to settle on. We didn't have much money, so we built very simple huts, with roofs that hardly protected us from rain.

2015

Over 20 years later, we still live on just 125 hectares of land. The average land entitlement in our region is 1.5 hectares per family.

Awra Amba has now grown to 500 inhabitants, but we are sharing a land big enough for just 11 families.

TEXTILE SHOP

Our biggest business is textiles, designed and produced in our village. We sell them in our own shop, as well as in many nearby villages and towns. The bestsellers are scarves and blankets, but we also make household textiles, clothes and bags.

RESTAURANT

Every day we receive visitors, be it backpackers from Europe or local traders. We cater for them in our restaurant, where we serve sweet tea, cold drinks, home-baked bread and traditional Ethiopian dishes.

TRADER'S MARKET

Recently, we have set up a trader's market in our village, where many local tradesmen come to sell and buy goods and food products.

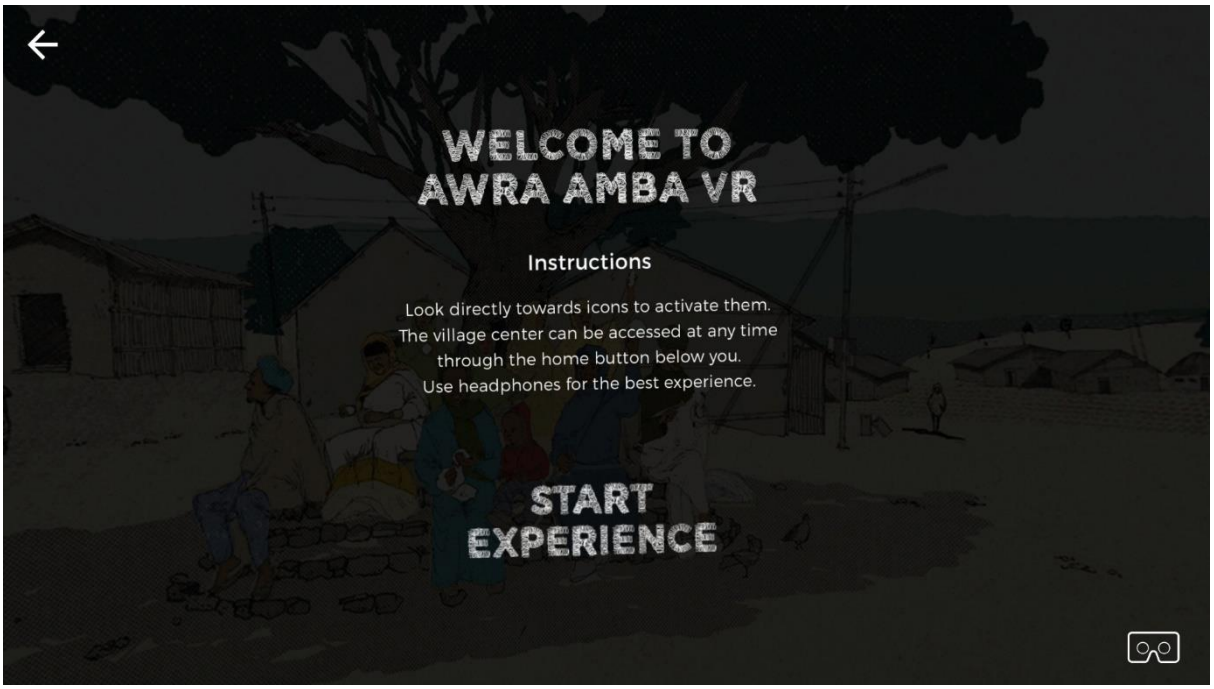
MILL

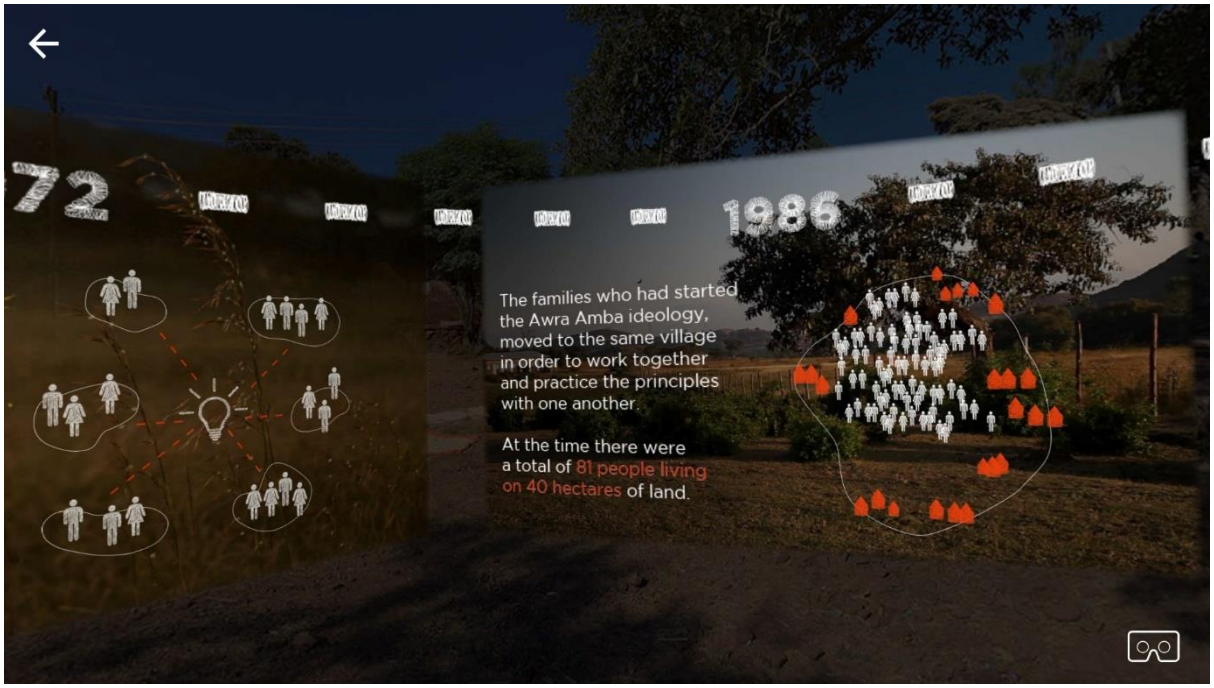
Our mill attracts thousands of customers from nearby towns and villages, who come with their donkeys, carrying sacks of grain to be milled. The milling process can take several hours, so people often visit our restaurant.

SUPERMARKET

With the influx of local tradesmen and students coming to the village daily, there was a need to open a supermarket that sells everyday commodities, normally only available in bigger towns. We stock up on everything from rubber boots to flashlights.

Appendix 3


WELCOME TO
AWRA AMBA VR
Instructions
Look directly towards icons to activate them.
The village center can be accessed at any time
through the home button below you.
Use headphones for the best experience.
START
EXPERIENCE


72
1986
The families who had started
the Awra Amba ideology,
moved to the same village
in order to work together
and practice the principles
with one another.
At the time there were
a total of 81 people living
on 40 hectares of land.

Appendix 4

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total paper - Total VR	.132	20	.200*	.961	20	.560

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Appendix 5

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Insertion Paper - VR	.226	20	.009	.913	20	.072

a. Lilliefors Significance Correction

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Omission Paper - VR	.227	20	.008	.933	20	.177

a. Lilliefors Significance Correction

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Substitution Paper - VR	.166	20	.153	.924	20	.118

a. Lilliefors Significance Correction

Appendix 6

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Paper - VR	.104	20	.200*	.954	20	.437

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Appendix 7

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Errors / words in paper - Errors / words in VR	.0066034537	.0379653531	.0084893110	-.011164879	.0243717859	.778	19	.446

Appendix 8

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Insertions / words paper - Insertions / words VR	-.002535709	.0055481699	.0012406085	-.005132333	.0000609143	-2.044	19	.055

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Substitutions / words paper - Substitutions / wordsVR	-.005043765	.0219383450	.0049055631	-.015311226	.0052236968	-1.028	19	.317

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Omissions - words paper - Omissions - words VR	.0018852498	.0283386536	.0063367156	-.011377648	.0151481480	.298	19	.769

Appendix 9

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Words per minute on paper - Words per minute in VR	5.500027531	12.72323175	2.845001109	-.454628225	11.45468329	1.933	19	.068

Appendix 10

Test Statistics[a]

	Expression and Volume Paper - Expression and Volume VR	Phrasing Paper - Phrasing VR	Smoothness Paper - Smoothness VR	Pace paper - Pace VR
Z	-1.732[b]	-.378[c]	-1.000[b]	-1.414[b]
Asymp. Sig. (2-tailed)	.083	.705	.317	.157

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.