A Randomized Control Trial of Computer-Based vs. Paper ABRACADABRA: Impacts on Reading and Spelling, Reading Motivation, Self-Esteem and Enjoyment

Kristen Sha

Department of Educational and Counselling Psychology

McGill University, Montreal

August 2018

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Master of Arts in Educational Psychology

Abstract

This pre-test post-test randomized control trial investigated the effectiveness of a web-based early literacy tool, A Balanced Approach for Children Designed to Achieve Best Results for All (ABRACADABRA), comparing the effects of delivery format (computer versus paper) on students' reading and spelling skills, reading motivation, self-esteem and enjoyment. Based on critiques of technology by Clark (1983) and the time-displacement hypothesis of technology (Vandewater, Bickham, & Lee, 2006), technology is predicted to have negative effects on learning and related percepts. Based on these models, hypotheses predicted negative effects of computer over paper ABRACADABRA delivery medium on participants' reading and spelling skills, reading motivation, reading self-esteem and reading enjoyment at post-test. The ABRACADABRA intervention involved delivering three weekly 15-minute supplemental reading sessions for eight weeks during the school day. Results showed comparable positive effects of both the computer-based and paper ABRACADABRA instruction on participants' reading and spelling at post-test and little evidence of difference by medium of intervention delivery on reading motivation, self-esteem, and enjoyment. It was concluded that the computer-based intervention does not have negative effects over its paper counterpart on students' literacy skills, and related literacy percepts, and provide no support for the time-displacement hypothesis in this context.

Keywords: ABRACADABRA, randomized control trial, reading, spelling, reading motivation, reading self-esteem, reading enjoyment, elementary school students

Résumé

Un essai randomisé et contrôlé pré-test/poste-test a été effectué afin d'analyser l'efficacité d'un outil d'alphabétisation précoce à base Web appelé ABRACADABRA. Cet outil compare les effets des formats de livraison (à l'ordinateur ou à la main) sur les compétences en lecture et en l'écriture des étudiants, leur motivation à la lecture, leur estime de soi, et leur plaisir à lire. Selon la critique des technologies de Clark (1983) et de l'hypothèse du déplacement temporel au sujet de la technologie (Vandewater, Bickham, & Lee, 2006), les technologies sont prévues d'avoir des effets négatifs sur l'apprentissage et les percepts liés à ces mécanismes. Selon ces modèles qui analysent la mise en œuvre de la ressource ABRACADABRA, en utilisant cette ressource, les hypothèses ont prévu des effets négatifs de l'apprentissage par l'ordinateur par rapport à la main en fonction des compétences en lecture et en orthographe des participants, la motivation à la lecture, l'estime de soi pendant la lecture, et le niveau du plaisir à lire à poste- test. L'étude a inclus la mise en œuvre de trois cours de lecture hebdomadaire supplémentaire de 15 minutes pour huit semaines pendant la journée scolaire. Les résultats indiquent des effets positifs comparable pour l'enseignement ABRACADABRA par ordinateur et par la main l'un et l'autre sur le niveau de la lecture et l'écriture des étudiants à poste-test. En outre, il n'y avait pas de différence de moyen de livraison sur la motivation à lire, l'estime de soi, ou le niveau du plaisir à lire. On a conclu que l'intervention à l'ordinateur ne possède pas des effets négatifs par rapport à l'intervention à la main sur les compétences d'alphabétisation des étudiants et des percepts liés, et ne supporte pas l'hypothèse du déplacement temporel dans ce cas.

Mots-clés : ABRACADABRA, essai randomisé et contrôlé, lecture, l'orthographe, motivation à la lecture, l'estime de soi en faisant la lecture, le plaisir à faire la lecture, étudiants du primaire

Acknowledgements

I would like to thank my supervisors, Professor Robert Savage and Professor Tara Flanagan, for their guidance, advice and support in the completion of this thesis. I am very fortunate to have supervisors who are always ready and willing to chat with me about my work, and to provide me with the necessary resources in order to complete this work. Apart from this thesis, Rob and Tara have also provided me with great mentorship throughout my time at McGill, and have fully supported me in my pursuit of further studies beyond the master's level. I am extremely grateful to have worked with them during my time at McGill, and I have certainly enriched my knowledge and understanding of human development, both in the academic realm and beyond.

I would like to express my deepest gratitude to my parents and my older sister Winnie, for their unconditional love and support in my academic pursuits. They are always standing behind me, ready to lend a listening ear or to offer kind words of support. Without their support, I would not have been able to complete my graduate studies and this thesis as smoothly and successfully as I have. I am truly indebted to them.

My community at McGill has been a core part of my wonderful learning experience during these two years. I would like to mention my Human Development MA & PhD cohort, ECP graduate student friends, Inter-Varsity friends from International Student Ministry and McGill Christian Fellowship, as well as my 3704 Peel housemates. These individuals have all played very special roles in my time at McGill, offering friendship, advice, comfort and support.

Table of Contents

List of Tables

List of Figures

A Randomized Control Trial of Computer-Based vs. Paper ABRACADABRA: Impacts on Reading and Spelling, Reading Motivation, Self-Esteem and Enjoyment

Does the medium of instruction influence learning? This question has been extensively debated within the realm of educational technology research. An extended discourse that has become known as the Clark-Kozma technology/media and learning debate (1983, 1994) explored contentions of pro- and anti-technology-based learning theorists on the effectiveness of technology as a medium of instruction over traditional methods (Bavelier, Green, & Dye, 2010). More specifically focused on literacy technologies for young children, previous investigation of ABRACADABRA (A Balanced Approach for Children Designed to Achieve Best Results for All), an evidence-based early literacy tool, contrasted the mode of delivery (computer versus paper) in teaching reading to lower elementary school students in the United Kingdom while keeping the content of the two interventions consistent. (McNally, Ruiz-Valenzuela, & Rolfe, 2016). McNally et al. found both forms of intervention to be effective compared to an untaught control group. McNally et al. also found the paper delivery to be somewhat more effective than the computer (McNally, Ruiz-Valenzuela, & Rolfe, 2016) on some measures, and analyses of relative effect sizes for reading. Following McNally et al.'s (2016) study, this study replicates the investigation on Grade 1 and Grade 2 Montreal students' reading skills, and expands it to explore spelling, reading motivation, reading self-esteem and reading enjoyment.

In this literature review, effective early reading interventions will be summarized. Educational technologies in reading and writing instruction will then be reviewed. The role of computer and paper programs in shaping elementary students' development in reading motivation, reading self-esteem and reading enjoyment will be discussed. An overview of online computer-based reading programs will follow. Afterwards there will be an overview of

ABRACADABRA. Finally, the current study will be described to explain how it addresses the effects of ABRACADABRA on an additional literacy skill (spelling) and its effects on reading motivation, reading self-concept and reading enjoyment.

Reading Instruction and Effective Early Interventions

Savage and Cloutier (2017) reviewed effective early reading interventions for several reading components: phonics, reading comprehension and fluency. Overview of these components will provide an overall idea of what works in reading instruction.

For phonics, previous meta-analyses demonstrate robust support for phonics instruction facilitating reading growth (Savage & Cloutier, 2017). Phonics instruction is most effective when it involves letter-sound and phonological awareness training together (Savage & Cloutier, 2017). Savage and Cloutier (2017) note, however, the need for more intervention studies to systematically control for length, time and content in order to ensure treatment quality.

Reading comprehension involves instruction of key language sub-components, for example vocabulary and morphological awareness, which can improve reading comprehension (Savage & Cloutier, 2017). Although there has been more research on morphology, still more work is required to understand the interactions between morphology, reading comprehension and reading accuracy (Savage & Cloutier, 2017).

As for fluency, repeated reading interventions generally lead to improved reading fluency in reading familiar, practiced passages (Savage & Cloutier, 2017). A limitation in this area of research, however, is the lack of well-designed experimental studies to provide further insight on reading fluency (Savage & Cloutier, 2017).

Over the years, there has been a substantial amount of research on effective early reading interventions. We now know what generally works in teaching phonics, reading comprehension

and fluency, although there is still room for improvements to be made in raising intervention quality and improving empirical rigor. One area that appears to be lacking in intervention studies is the number of studies directly contrasting medium of intervention (e.g. computer and paper) using equivalent reading content. This therefore leads to a review of educational technologies in reading instruction in order to examine the efficacy of such technologies in facilitating students' reading growth.

Reading Instruction and Educational Technologies

There is significant debate about the utility of educational technologies. A key question in such debates concerns whether the medium of delivery impacts the effectiveness of instruction. This question is captured in the classic Clark-Kozma (1983, 1994) technology/media and learning debate. Clark (1983) claimed that media presents a way of instruction but in and of itself does not influence student learning. His argument focuses on the hypothesis that the role of media and instructional methods have been mixed up – that instructional methods facilitate learning. He specifically highlights that different teaching methods can essentially be made into media formats, which leads him to question what are the unique contributions that media bring. Clark closes his argument with the need to examine whether media attributes have a special cognitive effect on learning, or if not, whether these attributes contribute meaningfully to other aspects crucial to learning.

In response, Kozma (1994) emphasized the need to explore the relationship between media and learning beyond a behaviourist (stimulus-response) framework. He pointed out the importance of looking at other key components of learning (e.g. cognitive/social/emotional), in particular how the fundamental structure and functions of media might influence these processes. Kozma postulates that learning involves rich, complex interactions between the learner and the

environment, and is not the result of passive reception to instructional delivery. He concludes with suggestions to investigate the role and efficacy of media in learning by employing theories, frameworks and interventions that allow for effects of media on cognitive and social processes to be explored.

The Clark-Kozma (1983, 1994) debate was situated in a time where educational technologies started to be used in classrooms. The use of technology still remains popular today. From the early 1980s until now, it is therefore important to review whether Clark and Kozma's arguments on the efficacy of educational technologies have since been addressed in research. Media-based educational technologies have been used in reading instruction as early as 1980 (Cheung & Slavin, 2012). Today these educational technologies are a widespread type of literacy (reading and writing) instructional method (Selwyn, 2012). This on-going trend in reading technology education therefore calls for an investigation on the effects of medium of instruction. Specifically, it will focus on comparing contemporary technology approaches and traditional paper methods in elementary literacy instruction. Of particular interest is whether each medium has unique contributions to learning.

Educational technology is broadly defined as "a variety of electronic tools and applications that help deliver learning materials and support learning process in classrooms" (Cheung & Slavin, 2012, p.201). These instructional mediums can include computer-assisted instruction, integrated learning systems and the use of multimedia such as videos (Cheung & Slavin, 2012). Computer-based reading programs are effective interventions in improving the reading outcomes of students with general reading abilities (Cheung & Slavin, 2012) and students who are struggling readers (Cheung & Slavin, 2013). Since the early 1990s, many researchers have conducted large-scale meta-analyses on the impact of educational technology in

reading instruction to investigate their impact on literacy (Becker, 1992; Blok, Oostdam, Otter, & Overmatt, 2002; Cheung & Slavin, 2012; Fletcher-Finn & Gravatt, 1995; Kulik & Kulik, 1991; Kulik, 2003; Ouyang, 1993; Soe, Koki, & Chang, 2000). All of the aforementioned reviews have found small to medium effects, ranging from +0.06 to 0.43, of the use of educational technology on reading performance (Cheung & Slavin, 2012).

In considering the small overall impact of educational technology on reading instruction, Savage and Cloutier (2017) draw attention to three methodological issues within such reviews: implementation of studies, quality of technology used in studies and theoretical and pedagogical alignment of technologies and their application in classrooms (Savage et al., 2013). Firstly, when teachers are provided with appropriate training on the use of technology and have on-going access to pedagogical and technological support and where study treatment integrity is high, the effectiveness of educational technology in reading instruction can reach effect sizes of 0.60 (Archer et al., 2014; Savage & Cloutier, 2017). Secondly, of the commercially available early literacy programs, only 15% of such programs provided targeted instruction for synthetic phonics and none provide instruction in other key reading skills such as phoneme segmentation or print-based knowledge (Grant et al., 2017). Thirdly, there is a gap between the theory and application of reading technologies, such that the majority of reading programs are not based upon, or incorporate theories of technological literacy (Savage & Cloutier, 2017). Therefore, incorporating these recommendations regarding literacy program design (theory and content) and literacy program use in the classroom in future research and future applied practice could yield more accurate and more reliably effective reading technology outcomes.

Furthermore, the question of what causes these modest positive effects of educational technology should be considered. In order to consider this issue, empirical investigation of

students' learning experience where content is held constant and the only difference lies in medium of instruction is needed. This is so that any differences in students' learning outcomes can be more reliably connected to medium of instruction, and from there investigation of unique contributions of medium on learning can be undertaken. Past meta-analyses have mostly focused on the efficacy of computer-assisted instruction (Blok et al., 2002; Fletcher-Finn & Gravatt, 1995; Ouyang, 1993; Soe et al., 2000), while others have looked at computer-based instruction (Kulik & Kulik, 1991), integrated learning systems (Becker, 1992) or a comprehensive overview of the aforementioned types of technologies (Cheung & Slavin, 2012). These different foci of approach means that a range of educational technologies with different designs, structures and content were investigated. None of these works, thus far, have specifically examined educational technologies where content is identical but the medium of instruction is different (e.g. computer vs. paper). Without paper-based technologies serving as an empirical comparison, this allows for limited scope of investigation on the deeper how's and why's of the effects of educational technologies in reading instruction. This serves as a key limitation in defining the roles and values of computer versus paper approaches in reading. Future research should therefore include direct comparisons of these approaches in order to investigate the unique contributions of each instructional medium. This work is thus undertaken here.

Writing Instruction: Technology vs. Pen/Pencil

There exists a modest literature of studies directly comparing the effects of digital and pen/pencil writing instruction on young children and elementary school students' writing abilities. For example, Berninger, Abbott, Augsburger, & Garcia, (2009) report that first-and third-grade students with and without learning disabilities showed a superior level of writing, in terms of writing length and writing speed, using handwriting compared to keyboarding.

Preschool children who received handwriting training demonstrated better outcomes in word writing and word reading compared to peers who experienced typing training (Kiefer et al., 2015). Third-grade students wrote at quicker speeds on a digital test, irrespective of the writing instruction tool (digital vs. pen/pencil) (Wollscheid, Sjaastad, Tomte, & Lover, 2016). Second-grade students demonstrated high rates of orthographic learning, with no significant differences between students who underwent printing or keyboarding practice (Ouellette & Tims, 2014). Of these studies, those by Oullette and Tims (2014) and Wolloscheid et al. (2016) demonstrate comparable quality in design, participants and procedure. Each incorporated closely comparable, rigorous experimental designs (either an experimental or quasi-experimental design). Learning content was effectively controlled and held constant between two experimental groups (computer versus paper). Participants of similar ages (Grade 2 to 3) in similar, modest sample sizes (between 40 to 60) participated in these interventions. Given the empirical similarities between these studies, the respective findings favouring digital writing tools or no substantial differences between digital or writing tools warrants a closer examination of the possible mechanisms behind these mixed findings. This work is thus undertaken here.

More generally, in a review of studies comparing the relative impacts of traditional and digital writing tools, Wollscheid, Sjaastad and Tomte (2016) highlight three key factors regarding the mixed findings for these two tools. First, studies based on different theoretical perspectives (e.g. cognitive-theoretical, neuroscience and learning, socio-cultural) reveal different findings and therefore do not allow for a general evaluation of whether traditional or digital writing tools are preferable. Second, the different theoretical perspectives within studies provide consistent supporting evidence for traditional writing tools (cognitive-theoretical perspective, neuroscience and learning – quasi-experimental/cohort designs) or digital writing

tools (socio-cultural perspective – qualitative designs) due to studies strictly adhering to the research design under their respective theoretical frameworks. Third, depending on the year in which studies were conducted, findings either showed preference for traditional writing tools (found in earlier studies when traditional writing methods were more common) or digital writing tools (found in later studies when digital writing methods are increasingly used). Taken together, these findings reflect the need for future research to utilize theoretical frameworks and research methodologies that are more open-ended and can fit with both technology or pen/pencil methods, and for there to be contemporary comparisons of both methods (e.g. how technology and pen/pencil methods are used today), so that a more comprehensive understanding of their educational value and effectiveness in today's classrooms can be garnered.

Reading Motivation

Reading motivation refers to an individual's beliefs, values and goals with respect to the concepts, processes and experiences of reading (Guthrie & Wigfield, 2000). This definition also encompasses people's behaviours that are related to reading (Aydemir & Ozturk, 2012). In motivational theory, motivation is categorized and differentiated by intrinsic motivation and extrinsic motivation, where the former is motivation that comes from within and the latter is motivation that comes from outside (Deci, Koestner, & Ryan, 1999). In the context of reading, these motivations are termed intrinsic reading motivation and extrinsic reading motivation.

Intrinsic reading motivation involves elements such as enjoying reading as a positive experience, valuing books as something enjoyable, attaching personal importance to reading and having interest in the topic contained in the reading material (Becker, McElvany, & Kortenbruck, 2010). Intrinsic reading motivation is attached to a learning goal orientation (Hellmich & Hoya, 2017). Extrinsic reading motivation includes receiving recognition, rewards or incentives from

others (Wang & Guthrie, 2004), meeting others' expectations or avoiding negative consequences (Hidi, 2000). Extrinsic reading motivation is attached to a performance goal orientation (Hellmich & Hoya, 2017).

Research on children's motivation towards engaging in reading related activities on the computer versus on paper reveals mixed findings. In favour of the computer, a study of fourth-grade students had higher desire to read when they read long and complex texts electronically versus in print (Greenlee-Moore & Smith, 1996). In support of paper texts, another study of fifth-grade students had significantly higher reading motivation when they read narrative and expository texts in print compared to the screen (Aydemir & Oeztuerk, 2012). Other studies discovered no significant differences in reading motivation between the two mediums. First- to third-grade students did not demonstrate significant differences in reading motivation for non-fiction e-books or non-fiction trade books read with adult support; however, overall students held positive attitudes towards using e-books and showed a preference for reading e-books over regular trade books (Barnyak & McNelly, 2016). Third-grade students did not demonstrate substantial differences in their reading motivation for e-books or traditional print books; instead children's ability to relate to the contents of the book, for example the setting, characters and theme, was more important (Jones & Brown, 2011).

In studies where students' outcomes clearly favoured the computer (Greenlee-Moore & Smith, 1996) or paper (Aydemir & Oeztuerk, 2012) experimental pre- and post-test designs (with treatment and control groups) and standardized literacy and reading motivation measures were used. In studies with mixed findings regarding students' outcomes, mixed methods designs specifically a combination of standardized literacy measures and student self-report motivation questionnaires were employed (Barnyak & McNelly, 2016; Jones & Brown, 2011). Despite

inconclusive findings in standardized motivation measures, these studies discovered a student-reported preference for engaging with technology-based texts where they found a sense of reader-text 'relatability'. This suggests that some aspects of reading motivation may not be accurately captured in standardized motivation measures (e.g. reading motivation scales), but may come through in subjective measures of motivation (e.g. interviews/questionnaires). Across all these studies sample sizes were modest (averaging between 30 to 40 participants) and participants were mostly in similar grade levels (upper elementary) recruited from one or two classrooms within the same school.

Overall this literature reveals that effect sizes demonstrating efficacy of the computer or paper are modest, and generalizability of results to the wider population of elementary students should be carefully considered. Moreover, only two studies (Aydemir & Oeztuerk, 2012; Greenlee-Moore & Smith, 1996) provided equivalent reading content across computer and paper delivery, allowing for direct comparisons of computer versus paper delivery effects on reading motivation to be made. The lack of studies with these empirical components in place serves as a limitation in reading motivation research. Future research using experimental pre- and post-test designs with larger participant sample sizes and careful controlling of reading content across experimental conditions (computer versus paper) could allow for deeper investigation of unique contributions and processes by which computer and paper methods influence students' reading motivation. This work is undertaken here.

Reading Self-esteem

The multi-dimensional model of self-concept (Shavelson, Hubner, & Stanton, 1976) describes self-esteem as a global, hierarchical construct that has many sub-dimensions such as academic self-concept and non-academic self-concepts. In this model, global self-concept

otherwise known as self esteem, is at the highest point of the model, followed by academic self-concepts (verbal and mathematical), and non-academic self-concepts (social, emotional and physical). While global self-concept and self-esteem are similar in nature and thus can be used interchangeably (Marsh & Martin, 2011; Shavelson et al., 1976), academic self-concepts appear to be domain-specific where within-domain self-concepts (e.g. verbal/reading) have stronger bearing on reading achievement (Marsh, Trautwein, Ludtke, Koller, & Baumert, 2006; Moller, Retelsdorf, Koller, & Marsh, 2011).

Reading self-esteem refers to an individual's beliefs about their ability to complete tasks within the domain of reading (Bandura, 1977) and additionally encompasses behaviours, attitudes and motivation related to reading (Gross, 2004). Perceptions of difficulty of reading, reading competence and reading attitude are core sub-components of reading self-esteem (Chapman & Tunmer, 1995; 1997). In an investigation of reading difficulty, reading self-concept and reading attitudes among children aged 5 to 10 (Chapman & Tunmer, 1995) overall reading difficulty was negatively associated with reading self-concept and reading attitudes; these constructs also had strong associations with reading performance (Wigfield, Cambria, & Ho, 2012). Altogether, children's reading self-concept, reading attitudes and reading difficulty perceptions are important constructs that relate to children's reading engagement and reading achievement. Self-esteem level can influence an individual's participation in reading – for example, individuals with high reading self-esteem will be more confident in their ability to try reading new story genres, authors or reading materials, whereas individuals with low reading self-esteem are less likely to do so (Hisken, 2012). The decision to engage in or avoid reading related activity influences students' overall reading achievement (Baker & Wigfield, 1999; Schunk, 1991).

The reciprocal effects model (Marsh, 1990; Marsh & Craven, 2006) posits that self-concept and academic achievement have a two-way relationship. This means that self-concept influences academic achievement, and academic achievement also influences self-concept. There is high stability in the relationship between reading achievement and reading self-concept (Retelsdorf, Koller, & Moller, 2014). In this relationship, specifically, there is strong support pointing to reading achievement as a predictor of reading self-concept (Retelsdorf et al., 2014). For example in Chen and Savage's (2014) study, participants' reading self-concept in particular perceptions of difficulty with reading improved after participating in reading interventions that facilitated positive literacy growth.

Since reading achievement is a strong predictor of reading self-concept, there should be a focus on promoting children's reading development through instructional programs. Students' participation in reading programs allows them to improve their reading skills, which then facilitates positive changes in their reading self-esteem (Newlin, 2003). There are a few examples of effective reading programs that focus on building students' reading skills and self-esteem (Hisken, 2012) such as the Paws for Reading program (Newlin, 2003), mentoring programs (King, 2002; Fives et al., 2013) and birthday book talks (Norton & Anfin, 2003). These innovative programs use different instructional foci (e.g. reading to dogs in the Paws for Reading program), instructional arrangement (e.g. a mentoring relationship between an adult and a student) and instructional material (e.g. birthday books to promote celebration of individual identity and personality). Despite the success of these programs, the use of birthday books for instance, has limited adherence to theory-driven models of reading motivation (reciprocal effects model) where improvement in reading attainment is emphasized. Moreover, these models are not structured in ways that focus on explicit, systematic instruction of reading components

(Armbruster, 2010) and so it is imperative to turn the focus on innovative, theory-driven programs.

Preliminary searches of the literature suggest that there exists little to no research on whether innovative, theory-driven computer-based programs or its paper counterparts can improve students' reading self-esteem. This represents a literature gap since innovative computer-based reading programs show promise in raising student literacy outcomes (Cheung & Slavin, 2012). An extensive literature search was therefore undertaken in major empirical databases such as ERIC (EBSCO) and PsycINFO using these search terms: "Reading self-concept" OR "academic self-concept" OR "reading self-confidence" OR "reading self-esteem" OR "reading self-efficacy" AND "Educational technology" OR "computer based training" OR "computer assisted instruction" OR "computer assisted language learning" AND "reading achievement" OR "reading outcomes" OR "reading performance" OR "reading readiness" OR "reading ability" OR "reading" AND "elementary school students" OR "primary school students" OR "grade 1 students" OR "grade 2 students". In each of these search results, only 103 items and 4 items were returned from ERIC and PsycINFO, respectively. All search items were reviewed, but none specifically examined the relationship between computer-based or traditional paper reading programs and reading self-esteem. Given the literature gap and the promise of computer-based reading programs, the researcher included self-esteem as one of the key variables in this study.

Reading Enjoyment

Reading enjoyment stems from using a variety of combined motivations and strategies during reading (Newman, Wehlage, & Lamborn, 1992). Individuals who enjoy reading read for a variety of reasons, such as pleasure gained from reading and interest in a book or topic (Ciampa,

2014) and understand the books within their described contexts (Guthrie et al., 1996). Reading enjoyment is often related to reading engagement (Kucirkova, Littleton, & Cremin, 2017). Readers who enjoy reading will often read for reasons of enjoyment, interest and learning (Baker, Dreher, & Guthrie, 2000; Guthrie, Wigfield, Metsala, & Cox, 1999; Tracey & Morrow, 2006). These readers will engage in diverse literacy activities to add to their knowledge and to interact with peers (Guthrie et al., 1996).

Research contrasting the effects of electronic versus print reading materials on children's reading enjoyment using show an overall increase in children's enjoyment of reading after engaging with electronic texts. Pre-school aged students demonstrated higher levels of persistence with storybook reading tasks when they participated in an adult-facilitated e-storybook session compared to an adult-facilitated traditional storybook session (Moody, Justice, & Cabell, 2010). Grimshaw, Dungworth, McKnight, and Morris, (2007) report that first-grade students did not show substantial differences in their enjoyment of reading the electronic or printed version of two children's storybooks, although enjoyment of the text was greater with accompanying narration provided by the e-books. A study of fourth-grade students showed greater enjoyment reading e-books than paperback books, although the difference in enjoyment between the two mediums was not significant (Chaudhry, 2014).

In these studies, it appears that the narrative features of technology-based texts (e.g. animations, sound effects, reader-text interactions) are a key ingredient in facilitating greater reading enjoyment. These studies all incorporated rigorous experimental designs with treatment and control groups. Participant sample sizes, however, varied quite a bit with modest samples between 20 to 25 to larger samples of more than 100 participants. Some variation in medium of technology-based texts (e.g. CD-ROMs, e-books) and storybook content also existed across

studies, making these studies harder to evaluate. Within studies, though, reading content was equivalent and well controlled for between experimental conditions (computer versus paper), serving as a strength in providing direct comparisons of efficacy between computer and paper delivery. Taking all of these factors into account, of the three studies, only Moody et al.'s (2010) study reported significant effects. Therefore, the overall positive gains in reading enjoyment found in these studies have only modest robustness. In a meta-analytic review of various types of educational technologies and their effectiveness, Cheung and Slavin (2012) found innovative technology applications (online-based computer reading programs) carried the most potential for advancing students' literacy development these are thus considered below.

Online Computer-Based Programs in Literacy Instruction

Wood, Grant, Gottardo, Savage and Evans (2017) investigated the effectiveness of a host of commercially available online programs (e.g. ABRACADABRA, Starfall) and offline (CD) programs (e.g. Arthur's Kindergarten Learning System, Clifford Reading early literacy software programs) for young learners. Overall findings revealed that online programs provided more comprehensive instruction for a wider range of reading skills (e.g. alphabetic knowledge, phonological awareness, text comprehension and more). The identified strengths of these online computer-based reading programs included their comprehensive nature to teach a variety of reading skills and its ability to cater to students with diverse reading abilities. Online programs addressed most of the skills and sub-skills within the five core aspects of reading: phonemic awareness, phonics, fluency, vocabulary and comprehension (National Reading Panel, 2000). Wood, Grant, Gottardo, Savage, and Evans, (2017) conclude that the far-reaching breadth of instructional capacity within online programs provide a full and balanced literacy approach and

curriculum that would make literacy learning accessible for learners with diverse abilities (Wood, Grant, Gottardo, Savage, & Evans, 2017).

A Balanced Reading Approach for Children Always Designed to Achieve Best Results for All (ABRACADABRA) Web-based Literacy Program

ABRACADABRA is a free web-based literacy program for educators, students and parents. Developed by the Centre for the Study of Learning and Performance at Concordia University, it is designed as an early literacy tool for Grade 1 and 2 students. ABRACADABRA provides a well-balanced literacy curriculum covering all nine skills within the reading taxonomy. It is highly interactive in nature, containing a broad selection of engaging activities and digital stories for students. An additional unique feature of ABRACADABRA is that it has two different delivery modes (computer and paper) with the same curriculum content in both, so allows a contrast of the medium of delivery above and beyond the content of the programs. Moreover, the efficacy of ABRACADABRA's design and structure is empirically proven and evidence-based.

ABRACADABRA has received excellent evaluations for its program design, with high instructional quality ratings for various reading skills within its program (Wood et al., 2017). ABRACADABRA has been intensively researched through a series of randomized control trial and quasi-experimental studies with successful student literacy outcomes (Abrami, Borohkovski, & Lysenko, 2015) and brought to scale in global contexts such as Northern Australia (Wolgemuth et al., 2014; Wolgemuth et al., 2013), Kenya (Abrami, Wade, Lysenko, Marsh, & Gioko, 2016) and Hong Kong (Cheung, Mak, Abrami, Wade, & Lysenko, 2016). For these reasons, ABRACADABRA is selected as the literacy tool to be investigated in this study.

ABRACADABRA contains a total of 32 activities on phonological awareness, reading fluency, comprehension and writing, most with multiple progressive levels of task difficulty. It also contains 21 stories with a selection of engaging follow-up activities. ABRACADABRA has a built-in feedback and support system for learners. Activities include informative, child-friendly instructions and appropriate, multi-level prompts to support the child in giving the correct answer. Correct answers will be given a visual-auditory confirmatory response (e.g. "Cat. That's the word!"), followed by a positive comment (e.g. "Awesome job!") commending the child's knowledge and effort. Incorrect answers will be followed by a visual-auditory prompt (e.g. "Try again") aiding the child in correcting the answer. After two consecutive incorrect answers, the system will provide visual-auditory representations (e.g. showing and saying the word "cat") of the correct answer.

In the ABRACADABRA teachers' zone, there is a comprehensive collection of multimedia and print-based teaching resources for teachers. Teachers can also use the communication tool, training videos and assessment tools as professional development activities. In the parents' zone, there is information on the reading activities and their targeted skills. There is also a rich collection of fun activity ideas and useful tips for using ABRACADABRA in the home. In summary, ABRACADABRA has garnered strong empirical support as an effective, evidence-based literacy tool.

Arguably however, there remain some unanswered questions around the ABRCADABRA technology. Prominent amongst these concerns the active ingredient in ABRACADABRA. To what extent for example does the animated medium of instruction or the instructional content drive effects reported? McNally, Ruiz-Valenzuela and Rolfe (2016) addressed these questions in a randomized control trial study. Trained teaching assistants (*n* =

49) delivered these programs to typical Year 1 students (n = 2,241) in the United Kingdom. Schools (n = 51) were first randomly allocated as either 1) treatment school (computer or paper intervention) or 2) control school (regular instruction). In treatment schools, students were then randomly allocated to either 1) computer-based ABRACADABRA, 2) paper ABRACADABRA or 3) standard literacy instruction. Within the standard literacy instruction control group, there were two sub-groups: students in strictly control schools and students belonging to control groups within treatment schools. This design component was implemented to assess 'spillover effects' (from students in intervention groups to those in control groups) in treatment schools.

The intervention lasted from October 2014 to May 2015, for a total of 8 months. This intervention was fully registered as a trial at: http://www.controlled-trials.com/isrctn/submission/. Fifteen-minute intervention sessions were held four times per week. Teaching assistants delivered sessions to small groups of three to five students. Treatment integrity was ensured by teaching assistants completing tracking sheets on students' learning progress, regular contact between this study's project team and schools, and regular school visits made by the project team and/or external evaluators to observe teaching assistants' intervention delivery. Treatment integrity evaluations were high overall and assessed as equivalent across both conditions.

This study's implementation team provided quality training to teaching assistants for 1.5 days before the start of the study. On the first full day of training, topics such as the teaching of reading, reading interventions, treatment integrity and student-group teaching and management were covered. Interactive training on the use of these reading interventions followed. After the initial training session, teaching assistants practiced implementing ABRACADABRA sessions with children outside of the study. Two weeks later, at the half-day training event, teaching

assistants shared their experiences, discussed intervention implementation topics and had questions regarding intervention delivery answered. Teaching assistants further practiced working with ABRACADABRA for the remainder of the session.

Several key findings resulted from this study. Firstly, both ABRACADABRA interventions led to positive gains in literacy, although the paper condition (β = .23) was somewhat higher than the computer-based (β = .138) one. Secondly, students who received either medium of ABRACADABRA instruction made two to three months' worth of literacy gains compared to their peers who received standard literacy instruction, based on students' post-test results from the Progress in Reading Assessment (PIRA) test. Thirdly, the positive effects of the intervention were more significantly pronounced in students who were from low socio-economic backgrounds for both the computer-based (β = .368) and paper (β = .396) ABRACADABRA conditions, and also for those who were below average readers at pre-test for both the computer-based (β =.215) and paper (β = .230) conditions.

McNally et al. (2016) discussed several factors responsible for the positive impact findings for both ABRACADABRA interventions. Teaching assistants and students both showed active participation in interventions. Teaching assistants showed strong commitment and responsibility to make thorough instructional preparations. Teaching assistants delivered high quality interventions: they followed lesson plans and advanced to higher difficulty levels when students were ready. Therefore the overall high quality of intervention delivery maximizes the likelihood of the ABRACADABRA program positively impacting students' literacy.

McNally et al. (2016) also raised possible factors concerning why the paper intervention had a marginally larger effect on students' post-test literacy achievement compared to the computer-based intervention. From the process evaluation findings, teaching assistants felt that

the paper intervention could be more flexibly adapted to meet students with various ability levels. This finding was especially highlighted in weaker readers pre-intervention who later showed the biggest improvement post-intervention. An additional reason could be due to difficulties teaching assistants faced in keeping to session timings in the paper condition. Overrunning sessions may mean relatively greater ABRACADABRA exposure, which could involve learning more content. These differences could facilitate variance in the overall efficacy of the computer versus paper delivery of ABRACADABRA on students' literacy outcomes.

This study's overall findings highlight that ABRACADABRA's curriculum content is effectively designed, evidenced by both ABRACADABRA conditions facilitating positive literacy impacts above and beyond standard classroom literacy instruction. ABRACADABRA's curriculum efficacy directly translates into closely comparable outcomes for both interventions compared to controls. This suggests that curriculum content carried more weight in facilitating overall literacy growth compared to medium of instruction.

McNally et al. (2016) concluded that future research requires incorporation of additional data (e.g. teaching assistants' progress tracking sheets and survey responses, as well as other sources) in the data analysis process. Pace of lessons is a possible confound that must also be empirically considered and stringently controlled for. This is to reduce the variation in session timings across both interventions and to ensure that differential amount of material covered is not the reason influencing ABRACADABRA intervention outcomes.

Following McNally et al.'s (2016) study, Bloom (2016) published a newspaper article discussing the paper ABRACADABRA intervention facilitating higher literacy outcomes for students post-intervention. The article included opinions on this finding from other professionals in the literacy field. Jonathan Solity, an honorary research fellow at University College London,

suggested that work on the computer versus paper involve different modes of engagement: recognition vs. recall respectively. Solity stated that "On a keyboard, it's potentially easier than if it's on paper, and [with pencil and paper] you've actually got to write. And if you're writing, my guess is that you are spending longer on the task" (Bloom, 2016, p. 108). The crux of Solity's comments is that the shallow processing involved in computer-based work may lead to less significant literacy growth in contrast to deeper processing and engagement required in pencil and paper work.

Solity's comments are theoretically captured within a broader time-displacement hypothesis of technology, which states that assuming time spent on literacy is zero-sum and that technology lacks or has sub-optimal intrinsic value and may thus displace other activities of greater educational value (Vandewater et al., 2006; Weis & Cerankosky, 2010). Such hypotheses are consistent with views of other broader critics of educational technology (Bavelier et al., 2010; Clark 1983). Empirically Weis & Cerankosky (2010) provide evidence from an RCT with mediation analyses that the introduction of video games led directly to reduced academic and behavioural functioning in boys. By extension here, a modified time-displacement hypothesis might suggest that the 'game' aspects of ABRACADABRA may displace more valuable paper-based instructional experiences.

Solity's claims regarding the negative carryover effects of technology must therefore be carefully reviewed. In McNally et al.'s (2016) study, spelling was not specifically investigated and hence there was no data on the effects of ABRACADABRA on this literacy domain. Moreover, the paper ABRACADABRA intervention did not involve any pens or pencils. Therefore Solity's claims about potential negative effects of technology do not relate specifically

to McNally et al.'s (2016) study. Constructively reframing Solity's comments would be to explore whether general negative carryover effects of technology exist over paper interventions.

The Present Study: ABRACADABRA Computer-Based vs. Paper and Pencil

The current study thus aims to address whether the medium of ABRACADABRA instruction carries any differential effects on overall literacy skills and social-emotional literacy components. A key part of this study will explore whether the computer-based intervention has general negative carryover effects on spelling skills over its paper counterpart. This study will replicate and extend McNally et al.'s (2016) work. The scope of this study is extended to three areas: effects of ABRACADABRA interventions on spelling skills, exploring reading motivation, reading self-esteem and reading enjoyment (following the above review showing no or little clear evidence on effects of medium of delivery on literacy outcomes) and ensuring equal pace of both ABRACADABRA interventions. The rationale for this empirical extension is two-fold: to conduct a more comprehensive study of ABRACADABRA's two instructional mediums on literacy skills and to contribute to the existing strong ABRACADABRA empirical foundation.

The research questions explored in this study are: 1) What are the effects of the computer-based vs. paper ABRACADABRA instruction on students' reading skills (listening comprehension, letter reading and word reading) post-intervention, 2) What are the effects of the computer-based vs. paper ABRACADABRA instruction on students' spelling skills (letter writing and word spelling) post-intervention, and 3) What are the effects of computer-based vs. paper ABRACADABRA reading activities on students' reading motivation, reading self-esteem and reading enjoyment post-intervention.

Based on relevant studies previously reviewed in this study and the time-displacement hypothesis in relation to the aforementioned research questions, there likely will be some

differential effects between computer-based and paper ABRACADABRA on reading and spelling. Based on reading motivation theory, literacy attainment levels will likely influence levels of reading motivation, reading self-esteem and reading enjoyment.

The hypotheses investigated are as follows:

$H1_0$: Students' reading skills will not be differentially impacted (with no negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention.

$H1_a$: Students' reading skills will be differentially impacted (with a negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention.

$H2_0$: Students' spelling skills will not be impacted differently (with no negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention.

$H2_a$: Students' spelling skills will be impacted differently (with a negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention.

$H3_0$: Levels of reading motivation, reading self-esteem and reading enjoyment will not vary (increase/decrease) according to students' literacy attainment post-intervention.

$H3_a$: Levels of reading motivation, reading self-esteem and reading enjoyment will vary (increase/decrease) according to students' literacy attainment post-intervention.

Method

Design

An experimental research design using a pre-test post-test randomized controlled trial was applied to compare the effectiveness of a computer-based version of ABRACADABRA with its paper version on literacy outcomes, reading motivation, reading self-esteem and reading enjoyment. The independent variable in this study was the delivery type of ABRACADABRA (computer vs. paper). The dependent variables were the literacy outcomes (reading and spelling

attainment), reading motivation, reading self-esteem and reading enjoyment. The control variables that were kept constant throughout the experiment included the pace of lessons and treatment integrity. A power analysis, using the GPower statistical software (Faul, Erdfelder, Lang, & Buchner, 2007) was conducted on this study's design, with $N = 34$ and with alpha = 0.05 and assuming a correlation of $r = .7$ between reading measures. This design is powered at 0.8 to detect $d = 0.4$, which are small to medium effect sizes.

For participants in the first elementary school, the pre-test took place in mid-November 2017 and the post-test in early February 2018. For participants in the second elementary school, the pre-test took place in early January 2018 and the post-test in mid-April 2018. All measures were administered at pre-test and post-test.

Randomization of participants involved two parts: 1) generating matched within-class participant pairings and 2) randomly allocating each participant within pairings to either the computer-based or paper ABRACADABRA condition. This randomization process was undertaken to avoid teacher effects being a possible confound influencing participants' intervention outcomes.

Participants

Participants included Grade 1 and Grade 2 students ($N = 34$) from two elementary English language schools within the Sir Wilfrid Laurier School Board in Laval, Quebec. All elementary English language schools within the Sir Wilfrid Laurier School Board ($N = 27$) were initially contacted with invitations to participate in this study. School principals were provided with the research study proposal, a McGill ethics approval certificate and a Sir Wilfrid Laurier School Board ethics approval certificate. Of these schools, two schools showed great interest and

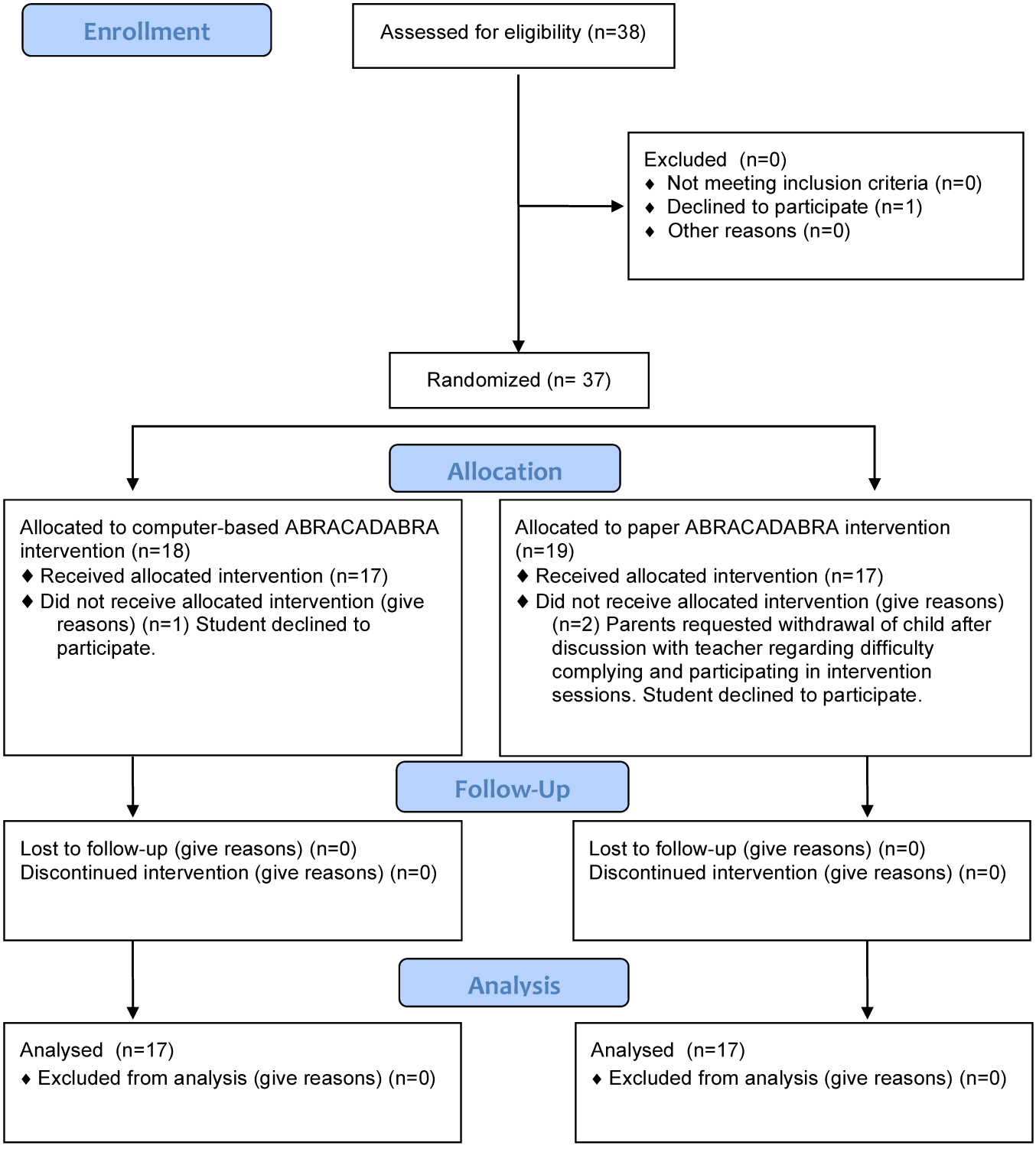
willingness to participate. Both schools have established working relationships with the ABRACADABRA research team from participating in previous ABRACADABRA studies.

The first school has actively participated in ABRACADABRA research projects conducted in previous years. Both the teaching staff and students have benefited from participating in these projects. For teachers, they were able to engage in professional development opportunities with regards to implementing ABRACADABRA in their daily instruction. For students, they showed positive gains in literacy outcomes. As a result of this, they were very eager to once again participate in an ABRACADABRA study.

The second school is relatively new in terms of their experience participating in ABRACADABRA research studies. At the first meeting with the principal and the Grade 2 teachers, the researcher explained the aim of the study to foster students' reading development. The teachers then relayed their concern for the reading skills of their weaker ability students. They responded favourably to having the study conducted at their school, especially since this reading intervention would benefit their struggling readers.

At the initial meetings with both schools, the researcher fully explained the study – aim, participants and procedure – and answered any queries or concerns from the principals or teachers. Principals were provided with the principal consent form, which was thoroughly read, understood and then signed. Teachers were provided with the teacher consent form, which was likewise read, understood and then signed. The researcher provided copies of the parental letter and consent form for participation in this study to teachers, which were then handed out to all students in their respective classes. Along with the parental consent forms, copies of parental questionnaires containing questions on children's background, demographics and language use were also distributed to students to bring home to their parents.

Figure 1. Consort Flow Diagram of Participants.



A sub-sample of Grade 1 students (*n* = 25, 12 boys, 13 girls) was from two Grade 1 classes in the first school. One Grade 1 male student withdrew from the study in week two, after his classroom teacher and parents had a discussion about his suitability to continue demonstrating a reasonable level of focus and compliance during reading sessions. Accordingly, this student's parent questionnaire data and pre-test data were omitted from further data analyses. Therefore, the final sub-sample of Grade 1 students was (*n* = 24, 11 boys, 13 girls). Participants' mean age was 6 years and 7 months (range = 6 years 2 months to 7 years 10 months).

The second sub-sample of students comprised of Grade 1 (*n* = 9, 7 boys, 2 girls) and Grade 2 students (*n* = 4, 1 boy, 3 girls) in the second school. These students were from one Grade 1 class and one Grade 2 class, respectively. Grade 1 participants' mean age was 6 years and 10 months (range = 6 years 3 months to 7 years 1 month). Grade 2 participants' mean age was 7 years and 9 months (range = 7 years 5 months to 8 years 2 months). During the pre-test process, one Grade 1 male student was unable to complete the pre-test measures due to behavioural difficulties. This student was therefore not included in the randomization process. Once the study began two Grade 1 male students declined to participate after attending the first session in week one. Accordingly, these three students' parent questionnaire data and pre-test data were omitted from further data analyses. Therefore, the final sub-sample of Grade 1 students was (*n* = 6, 4 boys, 2 girls). The number of participants in the Grade 2 sub-sample remains unchanged.

Overall 76% of parent questionnaires were returned. Students from both schools had typical reading and spelling. Students were unselected in that no exclusion criteria were applied in the selection of participants.

Apparatus and Materials

Wide Range Achievement Test (WRAT, Wilkinson & Robertson, 2006) Version 4 is an assessment that measures basic academic skills such as reading, spelling and math in individuals ages 5 to 94. The WRAT is norm-referenced with percentiles, standard scores, age equivalents and grade equivalents. The word reading and spelling sub-tests were individually administered to Grade 1 and Grade 2 students. It took approximately 15 to 20 minutes to complete. The word reading sub-test consists of two parts: letter reading (15 alphabet letters in total) and word reading (55 words in total). The Spearman Brown split half internal reliability in this sample was $\alpha = 0.93$. The spelling sub-test consists of two parts: letter writing (writing one's name followed by writing 13 alphabet letters) and word spelling (42 words in total). The Spearman Brown split half internal reliability in this sample was $\alpha = 0.92$.

Group Reading Assessment Diagnostic Evaluation (GRADE, Williams, 2001) - Levels 1-2 is a diagnostic reading assessment that measures the developmental skills of students from Pre-K to Grade 12. Measurement of students' reading skills includes the areas of: word reading, vocabulary, listening comprehension and reading comprehension. It determines what developmental skills have been mastered and where instruction or intervention is needed. The GRADE is a norm-referenced and developmentally based assessment tool that is typically administered in groups.

In this study GRADE Levels 1-2 will be used as it corresponds with the expected age and grade level of Grade 1 and Grade 2 students. Only the listening comprehension sub-test was administered because it is a brief measure compared to the other sub-tests (word reading, vocabulary and reading comprehension). This sub-test, together with the WRAT word reading sub-test, provides a broad understanding of students' receptive and productive language skills.

The listening comprehension sub-test consists of 17 questions where students listened to statements read aloud by the researcher and then responded by colouring in a small bubble for one of four picture options. It took approximately 15 to 20 minutes to complete. The Spearman Brown split half internal reliability in this sample was $\alpha = 0.63$.

Reading Self-Concept Scale (Chapman & Tunmer, 1995) is an assessment tool that measures the reading sub-component of academic self-concept. The tool includes three main areas of reading self-concept: 10 competence in reading items, 10 perceptions of difficulty with readings items and 10 attitudes towards reading items. Examples from each of these sub-scales include: "Can you work out hard words by yourself when you read?" (competence item), "Is reading to the class hard for you?" (difficulty item), and "Do you like word games in class?" (attitude item). The Reading Self-Concept Scale was administered to participants individually, whereby the researcher verbally delivered the questions, the participant provided verbal answers yes/no followed by a frequency rating on a Likert scale (1 – never and 5 – always), and then the research recorded the responses. It took approximately 15 minutes to complete. The Spearman Brown split half internal reliability in this sample for competence in reading items was $\alpha = 0.73$, perceptions of difficulty with reading items was $\alpha = 0.86$, and attitudes towards reading items was $\alpha = 0.63$.

Children's Author Recognition Test (CART, Senechal, LeFevre, Hudson, & Lawson, 1996) is a measure of children's reading engagement. This test is one way of measuring children's range of reading – that is, the higher number of children's authors that are known, the higher level of presumed reading engagement. The original list of children's authors included a total of 40 authors and 20 foil authors. The authors' names were extracted from 100 popular book titles taken from a variety of reliable sources (e.g. interviews with librarians, surveys of

children's literature bookstores, opinions of pre-schoolers' parents, bestseller lists from Western Publishers). The foil authors' names were obtained from the names of the *Developmental Psychology (Volume 17)* editorial board members.

Since this children's authors list was from 1996, the current study's researcher felt that it was necessary to update this list with authors that children today would be familiar with. In compiling the updated list, the researcher compared the list of Canadian children's authors from Senechal, LeFevre, Hudson and Lawson's (1996) paper with the 100 best all-time children's stories in Canada written by Valm (2017), a freelance writer and editor on topics such as children's literature, parenting and children. Valm's (2017) article on the 100 best all-time children's stories in Canada was based on children's literature experts' views of which stories fit the title of best all-time stories. In comparing these two sources, the researcher recorded the authors who appear in both lists. The rationale for this was to determine which children's authors have either written books that are still popular today or who continue to write popular, widely read stories, and therefore have time-honoured status and recognition in the field of children's literature. It was logically presumed that children today would likely know these authors based on their popular influence in children's literature. A total of 8 authors were extracted after careful cross-checking between Senechal et al.'s (1996) children's authors list and Valm's (2017) 100 best all-time children's stories in Canada.

The researcher also surveyed authors whose name appeared more than once in the 100 best all-time children's stories list, and authors who recently won children's literature awards in 2016. The greater number of times an author's name appeared within one list (e.g. 100 best all-time children's stories) or across a combination of any two or more lists (e.g. Senechal et al.'s (1996) authors list, 100 best all-time children's stories, children's 2016 literature award list and

children's 2015 literature award list), the higher the priority level for these authors to be included in the updated list. The rationale for this, once again, is that the more frequently an author appears in esteemed records pertaining to children's literature, the more popular, well-established and reliably recognizable these authors presumably are to children today. The children's 2016 and 2015 literature award lists were specifically included to provide time-appropriate context with regards to stories children would likely have read, so as to increase the likelihood that this list of authors' names would be contextually-relevant to participants. A total of 31 authors were extracted after thorough cross checking between Senechal et al.'s (1996) authors list, 100 best all-time children's stories and children's 2016 literature award list. Overall, a total of 39 authors were included from both cross checking stages.

During the pre-test process with the Grade 1 students in the first school, the researcher realized two errors had been made with the CART. The first error was an organizational and typography mistake where one of the author names had been repeated twice on the authors list. Hence, the original list had a total of 40 authors included, but upon detection of the duplicate author, the overall number of authors totalled 39. The second error concerned the list content where there was an omission of foil authors in the list of authors' names. This meant that all the authors on the CART were real authors, instead of a combination of real and foil authors.

Steps were taken to rectify the two errors in the CART. First, the duplicate author was removed from the authors list. Second, the researcher reviewed the total number of authors' names (real and foil) to be included in this corrected version by surveying Senechal et al.'s (1996) work and Allen, Cipielewski and Stanovich's (1992) work which both used the CART as one of the measures. The reason for surveying Allen et al.'s (1992) work in addition to Senechal et al.'s (1996) work was because of attention, logistical and time issues related to test

administration for the CART with Grade 1 and 2 participants. In Senechal et al.'s (1996) work, a total of 60 authors (40 real and 20 foil) were included, and the CART was administered to parents of children 3- to 6-years old. In the current study, the CART is to be administered to Grade 1 and 2 students, who have lower attention levels compared to adults. This point, in addition to the restrictions placed on total testing time (approximately 100 to 150 minutes for all five measures), means that logistically speaking and time-wise, it is not ideal to follow the test structure and content of the CART used in Senechal et al.'s (1996). In light of this, the researcher followed the CART test structure and content used in Allen et al.'s (1992), which was administered to fifth-grade students, who are closer in age and ability to the Grade 1 and 2 participants in the current study. The CART in Allen et al.'s (1992) study had 25 real authors and 15 foil authors, totalling 40 authors.

The corrected version of the CART therefore included 25 authors from the original 39, and another 15 foil authors to make a list of 40 authors. In selecting which of the 39 original authors to include in the corrected version, the researcher reviewed all existing authors and noted down authors who appeared in multiple sources (e.g. Senechal et al.'s (1996) authors list, 100 best all-time children's stories, children's 2016 literature award list and children's 2015 literature award list) or who had multiple entries within any of these sources. The researcher then calculated the average ranking of these authors in the different lists they appeared in. Authors with the highest rankings were kept on the list, while authors with the lowest rankings were removed from the list. Of the original 39 authors, 25 authors stayed on the list and the remaining 14 were removed. In compiling the names of foil authors, the researcher used an online random name generator website (http://random-name-generator.info/) to generate a total of 15 foil author names. After compiling the list of 25 real authors and 15 foil authors, the researcher used an

online random list generator (https://www.random.org/lists/) to create a list of authors wherein the placements of author names were fully randomized.

The Children's Author Recognition Test was administered individually to participants, whereby the researcher orally provided the authors' names, and the participant put either a check mark indicating "yes" in recognition of the author, or a cross mark indicating "no" in non-recognition of the author. It took approximately 15 to 20 minutes to complete.

With the Grade 1 students in the first school, the researcher had administered the erroneous version of the CART at pre-test. Due to the fact that these participants had already seen and completed the erroneous version at pre-test, the researcher did not re-administer the corrected version during the pre-test period to avoid participant bias from previous test exposure and possible test practice effects in a second test administration.

The researcher, however, did administer the corrected version to these Grade 1 participants at post-test, in order to obtain data that would allow for a comparison and analysis of pre- and post-test responses on these two different versions of the CART. With the Grade 1 and 2 students in the second school, the corrected version of the CART was administered at both pre- and post-test. The Spearman Brown split half internal reliability in this sample was ($\alpha = 0.96$).

Procedure

The reading intervention took place in the first half of the school year with the Grade 1 students in the first school, and then in the second half of the school year with the Grade 1 and Grade 2 students in the second school. The entire reading intervention including the pre- and post-test processes spanned mid November of 2017 to early April of 2018. With the Grade 1 students in the first school, the pre-test process began in early November. This was followed by an eight-week intervention that started in late November and ended in late January. Due to the

Christmas break, which lasted a week long, the intervention period was equally divided into two four-week blocks. The first four weeks of intervention were carried out before the Christmas break, and the last four weeks were carried out immediately after the Christmas break.

This intervention schedule arrangement was put in place with the rationale that two four-week intervention periods would allow students to receive an equivalent length and amount of reading instruction in the context of the whole ABRACADABRA intervention. This arrangement would reasonably buffer against possible negative impacts on students' learning brought about by the time gap in the intervention (e.g. students forgetting previously learned content regarding reading skills or stories, or students showing regressive use of previously acquired reading skills). Further, this arrangement was consistent across the computer-based and paper ABRACADABRA conditions, so no confounds (e.g. unequal exposure to interventions) that would influence the comparison between the effectiveness of the two conditions were introduced. The post-test process took place shortly afterwards in early February. There was a period of overlap during late January, where the intervention was ending for the Grade 1 students in the first school and the pre-test process had also been completed with the Grade 1 and Grade 2 students in the second school.

With the Grade 1 and Grade 2 students in the second school, the pre-test process began in early January. This was followed by an eight-week intervention that started in mid February and ended in early April. Due to the March break, which lasted a week, the intervention period was divided into two smaller blocks of three weeks and five weeks of intervention. The first three weeks of intervention were carried out before the March break, and the last five weeks were carried out immediately after the March break.

Unlike the intervention arrangement schedule for the first school, this time period in the academic year did not allow for the intervention process to be equally divided into two four-week intervention blocks. This specific arrangement of three- and five-week intervention blocks was consistent across the computer-based and paper ABRACADABRA, so no confounds (e.g. unequal exposure to interventions) that would influence the comparison between the effectiveness of the two conditions were introduced. This feature in the intervention schedule, with a one-week gap between the first and second intervention blocks, was kept consistent for the interventions held in the first and second school.

Randomization Process

Participants were randomly assigned to the experimental condition (computer-based ABRACADABRA) or the control condition (paper ABRACADABRA) using a matched pairs random assignment procedure. This process involved inspecting the initial assessments of students on their baseline reading (WRAT 4 word reading sub-test), spelling (WRAT 4 spelling sub-test) and broader language abilities (GRADE listening comprehension sub-test) and then sorting students into pairs based on similar ability levels on these tests. Matching of all pairs was undertaken within the same classrooms as this provides a control for overall classroom effects.

Participants were matched with another participant in the same class on their test scores from the WRAT 4 word reading and spelling sub-tests and the GRADE listening comprehension sub-test. Observable variables such as gender, ethnic or cultural background and socio-economic status were also considered as much as possible. Overall, 44% of participants were gender-matched, 13% were matched on ethnic or cultural background, and 13% of participants were matched on socio-economic status.

Within student pairs, participants were randomly assigned number 1 for the computer-based ABRACADABRA condition or number 2 for the paper version of ABRACADABRA using an online random number generator website (https://www.random.org/). For each randomized group, students were divided into small instructional groups consisting of three to four students.

In the computer-based ABRACADABRA condition, there were a total of six instructional groups. These groups consisted of three students, apart from one group of two students, due to there being only four students in the Grade 2 sub-sample. In the paper ABRACADABRA condition, there were also a total of six instructional groups. Four groups had three students, with the two remaining groups having two and four students, respectively. The slight variation in student numbers was due to an uneven number of Grade 1 students and the Grade 2 sub-sample only having four students. All instructional groups within each condition (computer vs. paper) had a balance of male and female students. The gender distribution was a 1:2 female to male ratio or a 1:2 male to female ratio for most instructional groups, apart from the two Grade 2 groups where the overall female to male student ratio was 3:1.

Reading Intervention

Prior to the intervention, all students provided verbal assent of their participation in the study. The reading intervention was conducted outside of the classroom in a quiet instructional space. Each group participated in the program for 15 minutes during a typical school day. Three sessions were held per week, for eight weeks, making a total of 24 sessions. For both groups, each session began with the experimenter defining one learning goal with each student. Then the experimenter instructed and supported the students in completing the learning activities, either

on the computer or on paper. The computer-based and paper ABRACADABRA interventions were identical, apart from the delivery format.

Shared Components of the Computer-based and the Paper ABRACADABRA Conditions

The curriculum was taken from the McNally et al. (2016) study. The curriculum structure, organization and content were identical across the computer-based and paper ABRACADABRA conditions. There were twenty weekly session plans altogether, with bi-weekly session plans for each children's story. Delivery format was the only difference between both conditions. For example, children's stories were either in a digital format (e.g. e-book) or a paper format (e.g. paper book). In this study, six out of twenty weeks' content was covered, equalling to 24 sessions carried out three times per week over a total of eight weeks (see Appendix A and B).

Daily session plans included a minimum of three to a maximum of five activities. These activities were either decoding (D), comprehension (C) or reading (R) tasks. These activities occurred in different combinations (e.g. DDRD, DDCD and CDRD). For each of these tasks, there was a specified length of time for completion. The total amount of time for completion of all these activities was fifteen minutes.

Students' reading progress was closely monitored via regular session tracking forms. When students responded correctly 80 to 90% of the time as a group to an activity after three consecutive entries, they had reached the mastery level for that level and then moved on to the next level.

The pace of all of the ABRACADABRA intervention sessions was carefully controlled for both conditions (computer and paper). Each session was 15 minutes in length. At the start of each session, the researcher recorded the start time on the student progress tracking form. At the

end of each session, the researcher recorded the end time on the same tracking form. During the intervention sessions, the researcher carefully timed each lesson component (e.g. alphabet, blending, tracking) so that sufficient time was set aside to effectively instruct and fully engage participants throughout the lesson. The aim of controlling the pace of lessons was to ensure that the delivery and instruction of the computer-based and paper ABRACADABRA interventions were equal in terms of amount of content covered, as this directly influences the effectiveness of the intervention and the quality of student learning.

Pace of lessons over the eight-week ABRACADABRA intervention by computer or paper condition was analyzed for each school to investigate whether lesson pacing was equivalent across both conditions. For the first school, there was no significant difference in overall pace of lessons between the computer condition (M = 15.79, SD = 1.14) and paper condition (M = 15.76, SD = 1.08), t (14) = 0.05, p = 0.97. For the second school, there was also no significant difference in overall pace of lessons between the computer condition (M = 16.00, SD = 1.11) and paper condition (M = 16.54, SD = 1.67), t (14) = -0.76, p = 0.46. This means that, overall, timing of lessons was effectively controlled for in the computer and paper conditions. This suggests that lessons were carried out for comparable lengths of time and equal amount of lesson content was covered in both conditions.

Computer-based ABRACADABRA Condition

In the computer-based ABRACADABRA condition, the researcher facilitated sessions using a Macbook laptop. The ABRACADABRA program is located at the website address: https://grover.concordia.ca/abra/en/. At the beginning of each session, the researcher set up the computer to display the ABRACADABRA program home page ready for the students to begin their reading session.

The computer-based ABRACADABRA curriculum contained twenty weekly session plans in total. The session plans were centred on a children's digital story. The digital stories were organized in a bi-weekly manner, where students spend two weeks reading and consolidating their knowledge and reading skills for each story. Bi-weekly session plans were structured in a way that students progressively build upon their storybook knowledge and reading skills through consistent exposure to a variety of reading activities.

Small groups of students took turns interacting with the ABRACADABRA program on the computer. Examples of students' interactions included answering a question, identifying a sound or reading a sentence. Every effort was made to ensure that all students in each group had equal opportunities to interact with the ABRACADABRA program. If there was not enough time for students to all have a turn at an activity, then the researcher ensured that they had a turn next session.

Paper ABRACADABRA Condition

In the paper ABRACADABRA condition, the researcher facilitated sessions using a range of instructional materials, according to the paper ABRACADABRA curriculum. Examples of such materials included: letter cards, mini whiteboards and stuffed toy animals. At the beginning of each session, the researcher set up the lesson's materials on a table ready for the students to begin their reading session.

The paper ABRACADABRA curriculum had twenty weekly session plans altogether. The session plans were focused on a children's story. The stories were arranged in a bi-weekly fashion, where students spend two weeks reading and reinforcing their knowledge and reading skills for each story. Bi-weekly session plans were organized in a way so that students gradually develop their storybook knowledge and reading skills through regular exposure to a range of

reading activities. Small groups of students engaged in paper ABRACADABRA activities. Examples of such activities included: filling out an "About Me" worksheet, playing an alphabet letters bingo game and participating in an auditory blending activity (e.g. listening to letter sounds and raising up the appropriate letter card for each sound).

Treatment Integrity

Treatment integrity (TI) refers to "the degree to which treatment is delivered as intended" (Yeaton & Sechrest, 1981, p.160). TI evaluation is crucial to intervention studies because it provides data on the internal validity of a study (e.g. that X caused Y) (Liaupsin, Ferro, & Umbreit, 2012). TI data allows for the attribution of observed effects to an intervention and the investigation of reasons for ineffective interventions – whether this ineffectiveness is due to the intervention itself, the delivery, or a combination of both factors (Liaupsin et al., 2012). Furthermore, TI data can be applied to examine how TI influences student outcomes (Capin, Walker, Vaughn, & Wanzek, 2017). In the context of reading, TI has been identified as a key predictor of student reading outcomes (O'Donnell, 2008). Hence in the current study, TI measures were implemented in order to assess the extent to which the ABRACADABRA interventions (specifically the delivery format: computer vs. paper) were responsible for students' reading and spelling outcomes, reading motivation, reading self-esteem and reading enjoyment levels.

In this study, the following four measures were put in place to ensure treatment integrity: 1) keeping to 15 minute session timings, 2) keeping records of student progress, 3) moving students to higher levels when they achieved 80% response accuracy, and 4) scheduling school visits to conduct treatment integrity observations. As stated previously, the pace of lessons component (keeping to 15 minute session timings) involved strictly timing lessons, closely

following lesson plans and accurately recording session start and end times. In monitoring student progress, the researcher kept daily session logs of students' performance on blending and decoding activities. When students, as a group, answered correctly 80% of the time, then the researcher moved on to a higher difficulty level. In scheduling the school-based treatment integrity observations, the researcher followed these steps. The researcher first selected either university undergraduate or graduate students studying psychology or education to act as external observers. Two students were then selected based on their field of study, availability and interest in participating in research/furthering their research experience. These two student observers were provided with an introduction to my study and training on how to observe and score the intervention sessions.

TI was evaluated using a TI observation checklist (McNally et al., 2016). The TI observation checklist included 9 areas of assessment: exposure, planning, instructional guidance, opportunities to succeed (levels/differentiation), group cohesion, pacing and efficient use of time, behaviour, adaptation/extension and overall rating and comments. Exposure refers to the amount of instruction and amount of engagement students have with the intervention activities. Planning is about how much preparation the intervention facilitator (in this case, the researcher) put in prior to the intervention session. Instructional guidance describes the quality and amount of facilitation provided for instructional activities. Opportunities to succeed (levels/differentiation) require the researcher to have an awareness of students' ability levels and demonstrate appropriate use of instructional activities to match students' abilities. Group cohesion speaks to group dynamics and how well students work together on instructional activities. Pacing and efficient use of time refers to whether the researcher planned the timing of the session and the instructional activities within the session accordingly. Behaviour is about how well students

behaved during the intervention session (e.g. the presence of off-task or disruptive behaviours) and whether the researcher used effective behavioural management techniques to address disruptive behaviours. Adaptation/extension describes whether the researcher delivered instructional activities with appropriate context or extension activities to reinforce the learning objectives of the session. Overall rating and comments encompasses all the TI observation components and requires a general evaluation of how well the session went.

The TI evaluation form provides a number rating from 0 to 4 (0 being the lowest and 4 being the highest) for each of the 9 categories. Each number rating corresponds to a descriptive quality rating (e.g. 0 – poor, 1 – barely adequate, 2 – adequate, 3 – good and 4 – excellent). At the end of the evaluation form, an overall score out of 36 is provided, along with a space for additional comments to be made.

TI observations were conducted on 20% of all intervention sessions. With a total of 24 intervention sessions, this meant that five intervention sessions were to be observed. In selecting the instructional groups for the TI observations, the researcher followed a randomization process with these steps. Using an online number generator website (https://www.random.org/), numbers 1 and 2 were used to determine which instructional group within which class of students would be observed for the computer-based and paper ABRACADABRA intervention sessions.

In selecting a student class for the computer-based ABRACADABRA observation session, number 1 indicated class 1 and number 2 indicated class 2. After selecting a student class, to determine the computer-based ABRACADABRA instructional group to be observed, number 1 indicated computer-based ABRACADABRA group 1 and number 2 indicated computer-based ABRACADABRA group 2.

The same procedure was repeated in the selection of a paper ABRACADABRA observation session. The only difference being that if one class had already been selected for the computer-based ABRACADABRA observation session, then by default the other class would be selected for the paper ABRACADABRA observation session.

For subsequent observation sessions, the student class not previously selected for either the computer-based or paper ABRACADABRA observation sessions were selected, with randomization then occurring within classes to determine the specific instructional group to be observed for each condition (computer vs. paper). This selection process ensured that instructional groups (computer vs. paper) from both classes had the opportunity to be selected for observation.

With both schools, the researcher fully informed the school principal and the teacher about the nature, purpose and arrangement of the treatment integrity observations and obtained their consent to conduct these observations in the school. Detailed parent letters describing these treatment integrity observations were also sent home to parents of students participating in the ABRACADABRA interventions.

Treatment integrity observation sessions were held on January 25th, 2018, February 1st, 2018, Thursday April 5th, 2018 and Thursday April 12th, 2018, respectively. A total of ten instructional groups were observed during these sessions. Sessions were held in the afternoon in the school library in the first school, and between late morning and noon in the school daycare in the second school.

For the first session, observer 1 was a fourth-year undergraduate psychology student and observer 2 was a master's counselling psychology student. For all subsequent sessions, observer 1 remained the same but observer 2 was a master's human development student. The reason for a

change in observer 2 between session 1 and later sessions was because the original observer 2 was unable to continue due to academic commitments.

During the observation sessions, the researcher delivered the ABRACADABRA interventions to instructional groups as per usual. Meanwhile, the two observers sat a short distance away from the researcher and the instructional group in order to observe and evaluate each of the sessions. The observers filled out the evaluation forms as the sessions were being delivered.

Inter-rater reliability (IRR) refers to the degree of agreement on observations between observers. IRR of TI observations were calculated using the formula: (number of agreements/number of observations) x 100. The product of this formula was multiplied by 100 in order to obtain a percentage of agreement. IRR was calculated for each school to evaluate whether treatment integrity was robust and comparable for both participant sub-samples. The overall IRR was 60% for all sessions in the first school. The overall IRR was 87% for all sessions in the second school. Typically, an acceptable standard of IRR is 70%. The second school meets and exceeds this IRR standard, whereas the first school falls a bit short. Comparing the two IRR percentages, school one is somewhat lower than school two. Part of this variance may have been due to a change in one of the observers between the first and second observation sessions at the first school, contributing to differences in perceptions and scoring of TI.

Further analyses on the nine treatment integrity components were conducted to investigate whether treatment integrity outcomes were equal across both conditions. For exposure, computer-based ABRACADABRA (*Mean rank* = 126.00) scored higher than the paper condition (*Mean rank* = 84.00), however this difference was not statistically significant $U = 48.00$, $p = 1.00$. For planning, computer-based ABRACADABRA (*Mean rank* = 126.00)

scored higher than the paper condition (*Mean rank* = 84.00), however this difference was not statistically significant $U = 48.00$, $p = 1.00$. For instructional guidance, computer-based ABRACADABRA (*Mean rank* = 133.50) scored higher than the paper condition (*Mean rank* = 76.50), however this difference was not statistically significant $U = 40.50$, $p = 0.35$. For opportunities to succeed, computer-based ABRACADABRA (*Mean rank* = 118.50) scored higher than the paper condition (*Mean rank* = 91.50), however this difference was not statistically significant $U = 40.50$, $p = 0.41$. For group cohesion, computer-based ABRACADABRA (*Mean rank* = 125.50) scored higher than the paper condition (*Mean rank* = 84.50), however this difference was not statistically significant $U = 47.50$, $p = 0.96$. For pacing and use of time, computer-based ABRACADABRA (*Mean rank* = 132.00) scored higher than the paper condition (*Mean rank* = 78.00), however this difference was not statistically significant $U = 42.00$, $p = 0.22$. For student behaviour, computer-based ABRACADABRA (*Mean rank* = 134.00) scored higher than the paper condition (*Mean rank* = 76.00), however this difference was not statistically significant $U = 40.00$, $p = 0.48$. For adaptation/extension, computer-based ABRACADABRA (*Mean rank* = 123.50) scored higher than the paper condition (*Mean rank* = 86.50), however this difference was not statistically significant $U = 45.50$, $p = 0.76$. For overall rating, computer-based ABRACADABRA (*Mean rank* = 120.00) scored higher than the paper condition (*Mean rank* = 90.00), however this difference was not statistically significant $U = 42.00$, $p = 0.51$. Overall, the computer-based condition scored slightly higher than the paper condition on all nine components. These differences, however, were all statistically non-significant. Differences may have been due to random selection of the additional two observation groups, which by chance resulted in computer groups both times, therefore leading to overall unequal numbers of computer ($N = 6$) and paper ($N = 4$) groups. This, in turn, may have

contributed to the computer-based ABRACADABRA groups consistently receiving slightly higher evaluation ratings. This means that the computer-based ABRACADABRA groups would naturally have higher ratings on the basis of higher number of observations, rather than purely because of the nature of intervention (computer vs. paper) being delivered and observed.

Results

Preliminary Data Analyses

Preliminary data analyses were used to check all data for normality, skewness and kurtosis. Based on descriptive analyses showing the full distribution of the data, it was confirmed there were no deviations from normality, no floor and ceiling effects and no significant outliers in the data.

Levene's test for homogeneity of variance was also explored for all data. For the pre-test data, GRADE – Listening Comprehension sub-test, F (1, 32) = 0.93, p = 0.34, WRAT – Word Reading sub-test, F (1, 32) = 0.10, p = 0.76, WRAT – Spelling sub-test, F (1, 32) = 0.00, p = 0.98, Children's Reading Self-Concept Scale, F (1, 32) = 0.34, p = 0.56, including the Difficulty sub-scale, F (1, 32) = 1.34, p = 0.26, Competence sub-scale, F (1, 32) = 0.02, p = 0.89, Attitude sub-scale, F (1, 32) = 1.02, p = 0.32, and the Children's Author Recognition Test, F (1, 32) = 0.64, p = 0.43.

For the post-test data, GRADE – Listening Comprehension sub-test, F (1, 32) = 0.41, p = 0.53, WRAT – Word Reading sub-test, F (1, 32) = 0.62, p = 0.44, WRAT – Spelling sub-test, F (1, 32) = 1.08, p = 0.31, Children's Reading Self-Concept Scale, F (1, 32) = 0.41, p = 0.53, including the Difficulty sub-scale, F (1, 32) = 3.38, p = 0.08, Competence sub-scale, F (1, 32) = 2.75, p = 0.11, Attitude sub-scale, F (1, 32) = 0.24, p = 0.63, and the Children's Author Recognition Test, F (1, 32) = 0.13, p = 0.72.

Levene's test for these measures was non-significant, such that the null hypothesis of equal variances was retained: the variances are equal. This means that there were no significant differences between the statistical distribution of participants' performance on these measures at pre-test or post-test.

Descriptive Data Analyses

Descriptive analyses of the data were conducted to evaluate participants' reading, spelling, reading motivation, reading self-concept and reading enjoyment outcomes, by ABRACADABRA intervention condition, between pre-test and post-test (see Table 1).

For computer-based ABRACADABRA participants, their pre- and post-test literacy scores are as follows: GRADE – Listening Comprehension sub-test (M = 13.59, SD = 3.10; M = 15.41, SD = 1.18), WRAT – Word Reading sub-test (M = 99.76, SD = 16.55; M = 111.00, SD = 15.68) and WRAT – Spelling sub-test (M = 104.18, SD = 17.29; M = 112.12, SD = 15.03). Their pre- and post-test scores relating to reading motivation, reading self-concept and reading enjoyment are as follows: Children's Reading Self-Concept Scale (M = 3.66, SD = 0.38; M = 3.61, SD = 0.45), including the Difficulty sub-scale (M = 2.70, SD = 0.87; M = 2.36, SD = 0.78), Competence sub-scale (M = 3.87, SD = 0.76; M = 4.01, SD = 0.78) and Attitude sub-scale (M = 4.42, SD = 0.55; M = 4.44, SD = 0.84) and the Children's Author Recognition Test (M = 14.06, SD = 13.64; M = 9.47, SD = 9.80).

For paper-based ABRACADABRA participants, their pre- and post-test literacy scores are as follows: GRADE – Listening Comprehension sub-test (M = 14.00, SD = 2.12; M = 15.53, SD = 1.18), WRAT – Word Reading sub-test (M = 96.29, SD = 17.81; M = 108.76, SD = 13.39) and WRAT – Spelling sub-test (M = 99.65, SD = 15.81; M = 110.06, SD = 12.07). Their pre- and post-test scores relating to reading motivation, reading self-concept and reading enjoyment are as

follows: Children's Reading Self-Concept Scale (*M* = 3.57, *SD* = 0.37; *M* = 3.46, *SD* = 0.37), including the Difficulty sub-scale (*M* = 2.82, *SD* = 0.68; *M* = 2.51, *SD* = 0.45), Competence sub-scale (*M* = 3.64, *SD* = 0.71; *M* = 3.72, *SD* = 0.55) and Attitude sub-scale (*M* = 4.24, *SD* = 0.58; *M* = 4.17, *SD* = 0.78) and the Children's Author Recognition Test (*M* = 12.82, *SD* = 11.98; *M* = 11.24, *SD* = 9.39).

Visual inspection of the means in Table 2 suggests substantial change in literacy outcomes between pre- and post-test that appears to be equivalent across both intervention conditions. There are slight decreases in participants' perceptions of difficulty associated with reading between pre- and post-test, which seem to be equivalent across both conditions. There are small increases in participants' reading competence self-ratings between pre- and post-test, which is just a bit higher in the computer group than the paper one. Participants' attitude towards reading showed a very small increase in the computer group between pre- and post-test, whereas those in the paper group showed a small decline. Lastly, there are small general declines in participants' reading enjoyment between pre- and post-test that seem to be slightly higher in the computer group than the paper one.

Table 1. Means, Standard Deviations and Cohen's *d* Effect Sizes for the Pre-test and Post-test Literacy Measures by Intervention Group.

	Computer-based ABRACADABRA			Paper ABRACADABRA		
Measure	Pretest	Posttest	Effect Size	Pretest	Posttest	Effect Size
GRADE – Listening Comprehension[a]	13.59 (3.10)	15.41 (1.18)	0.69	14.00 (2.12)	15.53 (1.18)	0.58
WRAT – Reading[b]	99.76 (16.55)	111.00 (15.68)	0.66	96.29 (17.81)	108.76 (13.39)	0.73
WRAT – Spelling[b]	104.18 (17.29)	112.12 (15.03)	0.48	99.65 (15.81)	110.06 (12.07)	0.63
Reading Self-Concept Scale[a]	3.66 (0.38)	3.61 (0.45)	-0.14	3.57 (0.37)	3.46 (0.37)	-0.30
RSCS Difficulty	2.70 (0.87)	2.36 (0.78)	-0.12	2.82 (0.68)	2.51 (0.45)	-0.11

RSCS Competence	3.87 (0.76)	4.01 (0.78)	0.04	3.64 (0.71)	3.72 (0.55)	0.02
RSCS Attitude	4.42 (0.55)	4.44 (0.84)	0.00	4.24 (0.58)	4.17 (0.78)	-0.02
Children's Author Recognition Test[a]	14.06 (13.64)	9.47 (9.80)	-0.36	12.82 (11.98)	11.24 (9.39)	-0.12

Note.
GRADE = Group Reading Assessment and Diagnostic Evaluation; WRAT = Wide Range Achievement Test IIII; RSCS Difficulty = Reading Self-Concept Difficulty Sub-scale; RSCS Competence = Reading Self-Concept Competence Sub-scale; RSCS Attitude = Reading Self-Concept Attitude Sub-scale; Recognition Test[a] = Children's Author Recognition Test
[a]Values are represented by raw scores. [b]Values are represented by standard scores.

Inferential Analyses

First, independent samples *t*-tests were carried out to analyze participants' pre-test scores. The dependent variables included participants' pre-test scores on the GRADE – Listening Comprehension sub-test, WRAT – Word Reading sub-test, WRAT – Spelling sub-test, Children's Reading Self-Concept Scale and Children's Author Recognition test. The independent variable was the ABRACADABRA intervention condition (computer vs. paper). Results revealed that the computer ABRACADABRA and paper ABRACADABRA groups were not significantly different from each other at pre-test on listening comprehension, t (32) = -0.45, p = 0.66, reading, t (32) = 0.59, p = 0.56, spelling, t (32) = 0.98, p = 0.43, reading self-concept, t (32) = 0.56, p = 0.46, and reading enjoyment, t (32) = 0.43, p = 0.78, which means that both groups were comparable on all these measures, and confirmed that the matching process undertaken prior to randomization was successful in controlling for selection bias in primary outcome measures.

Next, independent samples *t*-tests were also conducted on participants' parental questionnaire data. Variables explored include the gender ratio of participants, participants' chronological age (years), parent-reported learning difficulties of their child, participants' English reading frequency, participants' French reading frequency, mother's education level, mother's native language, mother-child home language, father's native language, and father-

child home language. Results showed that the computer ABRACADABRA and paper ABRACADABRA groups were not significantly different from each other with regards to gender ratio, t (32) = -0.67, p = 0.51, chronological age (years), t (24) = 1.00, p = 0.33, parent-reported learning difficulties – the t statistic could not be computed because both groups are 0, English reading frequency, t (24) = -0.82, p = 0.42, French reading frequency, t (22) = 0.38, p = 0.71, mother's education level, t (24) = -0.58, p = 0.57, mother's native language, t (21) = 0.11, p = 0.91, mother-child home language, t (22) = -0.88, p = 0.39, father's native language, t (21) = -0.86, p = 0.40, and father-child home language, t (22) = -1.52, p = 0.14, which indicates that both groups were comparable on all these variables, and that there was low risk of selection bias in this study. The means and *SD*s for the pre-test scores and participants' parental questionnaire data are shown in Table 2. Effect sizes for participants' mean scores across all measures (listening comprehension, reading, spelling, reading self-concept and reading enjoyment) were computed using Cohen's *d* equation. This equation reports the mean differences of each condition within total sample, presented as (post-test – pre-test)/(pooled sample pre-test *SD*) (see Table 2 for effect sizes).

Table 2. Matching Characteristics of the Intervention Sample by Condition.

Condition	Computer-based ABRACADABRA	Paper ABRACADABRA	Significance
Gender (% male)[a]	53	41	0.51 *ns*
Gender (% female)[a]	47	59	0.51 *ns*
Chronological age in years[a]	6.42	6.21	0.33 *ns*
Parent-reported learning difficulties[c]	0	0	0.00[d]
English reading frequency[a]	2.08 (0.95)	2.38 (0.96)	0.42 *ns*
French reading frequency[a]	3.09 (1.22)	2.92 (0.95)	0.71 *ns*
Mother's education[a]	3.15 (1.07)	3.38 (0.96)	0.57 *ns*
Mother's native	2.33 (1.07)	2.27 (1.55)	0.91 *ns*

language			
Mother-child language[a]	2.08 (0.90)	2.50 (1.38)	0.39 *ns*
Father's native language	2.17 (0.94)	2.64 (1.63)	0.40 *ns*
Father-child language[a]	1.83 (0.58)	2.33 (0.99)	0.14 *ns*
GRADE – Listening Comprehension[a]	13.59 (3.10)	14.00 (2.12)	0.66 *ns*
WRAT – Reading[b]	99.76 (16.55)	96.29 (17.81)	0.56 *ns*
WRAT – Spelling[b]	104.18 (17.29)	99.65 (15.81)	0.43 *ns*
Reading Self-Concept Scale[a]	3.66 (0.38)	3.57 (0.37)	0.46 *ns*
RSCS Difficulty	2.70 (0.87)	2.82 (0.68)	0.26 *ns*
RSCS Competence	3.87 (0.76)	3.64 (0.71)	0.89 *ns*
RSCS Attitude	4.42 (0.55)	4.24 (0.58)	0.32 *ns*
Recognition Test[a]	14.06 (13.64)	12.82 (11.98)	0.78 *ns*

Note.
GRADE = Group Reading Assessment and Diagnostic Evaluation; WRAT = Wide Range Achievement Test IIII; RSCS Difficulty = Reading Self-Concept Difficulty Sub-scale; RSCS Competence = Reading Self-Concept Competence Sub-scale; RSCS Attitude = Reading Self-Concept Attitude Sub-scale; Recognition Test[a] = Children's Author Recognition Test
[a]Values are represented by raw scores. [b]Values are represented by standard scores. [c]Values are represented by percentage. [d]Significance unable to be computed due to percentages being a value of 0.

A series of 2 Condition (computer-based ABRACADABRA vs. paper ABRACADABRA) x 2 Test (pre-test and post-test) mixed model ANOVAs were then conducted to compare participants' outcomes on a series of measures at pre-test and post-test. These analyses looked at the impact of intervention condition on listening comprehension, reading, spelling, reading self-concept (including reading motivation and reading self-esteem), and reading enjoyment.

Main Analyses

Listening comprehension.

There was a significant effect of Test (pre-test vs. post-test) on participants' listening comprehension skills at the $p < .05$ level for both the computer-based and paper

ABRACADABRA Conditions, $\Lambda = 0.62$, F (1, 32) = 19.77, $p < 0.05$. There was, however, no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' listening comprehension skills at the $p < .05$ level, $\Lambda = 1.00$, F (1, 32) = 0.15, $p = 0.70$. These results suggest there is substantial improvement in listening comprehension skills for participants in the computer-based and paper ABRACADABRA conditions between pre-test and post-test. There are, however, no significant differences in participants' listening comprehension skills, based on whether they received computer-based or paper ABRACADABRA instruction between pre-test and post-test.

A closer examination of the listening comprehension mean scores shows overall improvement for participants in both conditions. In the computer-based ABRACADABRA condition, participants' listening comprehension skills improved between pre-test ($M = 13.59$, $SD = 3.10$) and post-test ($M = 15.41$, $SD = 1.18$), $d = 0.69$, showing a medium effect size for improvement. In the paper ABRACADABRA condition, participants' listening comprehension skills similarly improved between pre-test ($M = 14.00$, $SD = 2.12$) and post-test ($M = 15.53$, $SD = 1.18$), $d = 0.58$, also showing a medium effect size for improvement.

Given that there is a non-significant p value for the Test x Condition interaction, this suggests that neither technology nor paper methods have significant differential impact over the alternate condition in terms of participants' listening comprehension skills between pre- and post-test. The comparable positive effect sizes for participants' listening comprehension in both the computer-based and paper ABRACADABRA groups indicates there are no negative effects of technology on participants' listening comprehension skills.

Reading.

There was a significant effect of Test (pre-test vs. post-test) on participants' reading skills at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.25$, $F(1, 32) = 92.91$, $p < 0.05$. There was, however, no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' reading skills at the $p < .05$ level, $\Lambda = 0.99$, $F(1, 32) = 0.25$, $p = 0.62$. These results suggest there is substantial improvement in reading skills for participants in the computer-based and paper ABRACADABRA conditions between pre-test and post-test. There are, however, no significant differences in participants' reading skills, based on whether they received computer-based or paper ABRACADABRA instruction between pre-test and post-test.

A closer examination of the reading mean scores shows overall improvement for participants in both conditions. In the computer-based ABRACADABRA condition, participants' reading skills improved between pre-test ($M = 99.76$, $SD = 16.55$) and post-test ($M = 111.00$, $SD = 15.68$), $d = 0.66$, showing a medium effect size for improvement. In the paper ABRACADABRA condition, participants' reading skills similarly improved between pre-test ($M = 96.29$, $SD = 17.81$) and post-test ($M = 108.76$, $SD = 13.39$), $d = 0.73$, showing a medium effect size for improvement.

Given that there is an non-significant p value for the Test x Condition interaction, this suggests that neither technology nor paper methods have significant impact over the alternate condition in terms of participants' reading skills between pre- and post-test. The comparable positive effect sizes for participants' reading in both the computer-based and paper

ABRACADABRA groups points to no negative effects of technology on participants' reading skills.

Spelling.

There was a significant effect of Test (pre-test vs. post-test) on participants' reading skills at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.37$, $F(1, 32) = 53.92$, $p < 0.05$. There was, however, no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' spelling skills at the $p < .05$ level, $\Lambda = 0.97$, $F(1, 32) = 0.98$, $p = 0.33$. These results suggest there is substantial improvement in spelling skills for participants in the computer-based and paper ABRACADABRA conditions between pre-test and post-test. There are, however, no significant differences in participants' spelling skills, based on whether they received computer-based or paper ABRACADABRA instruction between pre-test and post-test.

A closer examination of the spelling mean scores shows overall improvement for participants in both conditions. In the computer-based ABRACADABRA condition, participants' spelling skills improved between pre-test ($M = 104.18$, $SD = 17.29$) and post-test ($M = 112.12$, $SD = 15.03$), $d = 0.48$, showing a medium effect size for improvement. In the paper ABRACADABRA condition, participants' spelling skills similarly improved between pre-test ($M = 99.65$, $SD = 15.81$) and post-test ($M = 110.06$, $SD = 12.07$), $d = 0.63$, showing a medium effect size for improvement.

Given that there is a non-significant p value for the Test x Condition interaction, this suggests that neither technology nor paper methods have significant impact over the alternate condition in terms of participants' spelling skills between pre- and post-test. The comparable

positive effect sizes for participants' spelling in both the computer-based and paper ABRACADABRA groups highlights no negative effects of technology on participants' spelling skills.

Reading self-concept (reading motivation and reading self-esteem).

There was no significant effect of Test (pre-test vs. post-test) on participants' reading self-concept at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.97$, F (1, 32) = 0.87, $p = 0.36$. There was also no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' reading self-concept at the $p < .05$ level, $\Lambda = 1.00$, F (1, 32) = 0.71, $p = 0.79$. These results suggest that, generally, participants' reading self-concept did not differ significantly between pre-test and post-test. Moreover, participants' reading self-concept did not differ significantly based on which intervention condition (computer vs. paper ABRACADABRA) they were in between pre-test and post-test.

A closer examination of the reading self-concept mean scores shows a slight overall decrease in reading self-concept for participants in both conditions. In the computer-based ABRACADABRA condition, participants' reading self-concept dropped between pre-test (M = 3.66, SD = 0.38) and post-test (M = 3.61, SD = 0.45), d = -0.14, showing a very small effect size for decline. In the paper ABRACADABRA condition, participants' reading self-concept similarly dropped between pre-test (M = 3.57, SD = 0.37) and post-test (M = 3.46, SD = 0.37), d = -0.30, a small effect size for decline.

Given that there is a non-significant p value for the Test x Condition interaction and small or near zero effect sizes for participants' reading self-concept in both the computer-based and paper ABRACADABRA groups, this suggests that 1) technology does not negatively impact

participants' reading self-concept over paper methods over time and 2) the general decline in participants' reading self-concept is not statistically significant.

Difficulty sub-scale.

There was a significant effect of Test (pre-test vs. post-test) on participants' perceptions of difficulty associated with reading at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.83$, $F\ (1, 32) = 6.58$, $p < 0.05$. There was, however, no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' perceptions of difficulty at the $p < .05$ level, $\Lambda = 1.00$, $F\ (1, 32) = 0.01$, $p = 0.93$. These results suggest there are substantial differences in perceptions of difficulty associated with reading for participants in the computer-based and paper ABRACADABRA conditions between pre-test and post-test. There are, however, no significant differences in participants' perceptions of difficulty, based on whether they received computer-based or paper ABRACADABRA instruction between pre-test and post-test.

A closer examination of the perceptions of difficulty mean scores shows a slight overall decrease in these perceptions for participants in both conditions. In the computer-based ABRACADABRA condition, participants' perceptions of difficulty dropped between pre-test ($M = 2.70$, $SD = 0.87$) and post-test ($M = 2.36$, $SD = 0.78$), $d = -0.12$, showing a very small effect size for decline. In the paper ABRACADABRA condition, participants' perceptions of difficulty similarly dropped between pre-test ($M = 2.82$, $SD = 0.68$) and post-test ($M = 2.51$, $SD = 0.45$), $d = -0.11$, showing a very small effect size for decline.

Given that there is a non-significant p value for the Test x Condition interaction, this suggests that neither technology nor paper methods have significant impact over the alternate

condition in terms of participants' perceptions of difficulty associated with reading between pre- and post-test. The comparable small effect sizes for participants' perceptions of difficulty in both the computer-based and paper ABRACADABRA groups suggests participants find reading to be a little less difficult than before, which seem to be equivalent across both conditions.

Competence sub-scale.

There was no significant effect of Test (pre-test vs. post-test) on participants' reading competence self-ratings at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.98$, $F\ (1, 32) = 0.56$, $p = 0.46$. There was also no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' reading competence self-ratings at the $p < .05$ level, $\Lambda = 1.00$, $F\ (1, 32) = 0.05$, $p = 0.83$. These results suggest that, generally, participants' reading competence self-ratings did not differ significantly between pre-test and post-test. Moreover, participants' reading competence self-ratings did not differ significantly based on which intervention condition (computer vs. paper ABRACADABRA) they were in between pre-test and post-test.

A closer examination of the reading competence mean scores shows a very slight overall increase in these ratings for participants in both conditions. In the computer-based ABRACADABRA condition, participants' self-ratings rose between pre-test ($M = 3.87$, $SD = 0.76$) and post-test ($M = 4.01$, $SD = 0.78$), $d = 0.04$, showing a very small effect size for improvement. In the paper ABRACADABRA condition, participants' self-ratings similarly rose between pre-test ($M = 3.64$, $SD = 0.71$) and post-test ($M = 3.72$, $SD = 0.55$), $d = 0.02$, showing a very small effect size for improvement.

Given that there is a non-significant *p* value for the Test x Condition interaction and small or near zero effect sizes for participants' reading competence self-ratings in both the computer-based and paper ABRACADABRA groups, this suggests that 1) technology does not negatively impact participants' reading competence over paper methods over time and 2) there is a general improvement in participants' reading competence, however it is not statistically significant.

Attitude sub-scale.

There was no significant effect of Test (pre-test vs. post-test) on participants' attitude towards reading at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 1.00$, $F\ (1, 32) = 0.02$, $p = 0.89$. There was also no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' attitudes at the $p < .05$ level, $\Lambda = 1.00$, $F\ (1, 32) = 0.10$, $p = 0.76$. These results suggest that, generally, participants' attitude towards reading did not differ significantly between pre-test and post-test. Moreover, participants' reading attitudes did not differ significantly based on which intervention condition (computer vs. paper ABRACADABRA) they were in between pre-test and post-test.

A closer examination of the reading attitude mean scores shows a very slight increase in these ratings for participants in the computer-based ABRACADABRA condition. In the computer-based condition, participants' attitudes rose between pre-test ($M = 4.42$, $SD = 0.55$) and post-test ($M = 4.44$, $SD = 0.84$), $d = 0.00$, showing a zero effect size for improvement. In contrast, in the paper ABRACADABRA condition, participants' attitudes dropped between pre-test ($M = 4.24$, $SD = 0.58$) and post-test ($M = 4.18$, $SD = 0.78$), $d = -0.02$, showing a very small effect size for decline.

Given that there is a non-significant *p* value for the Test x Condition interaction and small or near zero effect sizes for participants' reading attitude self-ratings in both the computer-based and paper ABRACADABRA groups, this suggests that 1) technology does not negatively impact participants' reading attitudes over paper methods over time and 2) there is an extremely small improvement in participants' reading attitudes in the computer-based ABRACADABRA condition, however it is not statistically significant.

Reading enjoyment.

There was a significant effect of Test (pre-test vs. post-test) on participants' reading enjoyment at the $p < .05$ level for both the computer-based and paper ABRACADABRA Conditions, $\Lambda = 0.84$, $F\ (1, 32) = 6.33$, $p < 0.05$. There was, however, no significant interaction between intervention Condition (computer ABRACADABRA vs. paper ABRACADABRA) and Test (pre-test vs. post-test) on participants' reading enjoyment at the $p < .05$ level, $\Lambda = 0.96$, $F\ (1, 32) = 1.49$, $p = 0.23$. These results suggest there are substantial differences in reading enjoyment for participants in the computer-based and paper ABRACADABRA conditions between pre-test and post-test. There are, however, no significant differences in participants' reading enjoyment, based on whether they received computer-based or paper ABRACADABRA instruction between pre-test and post-test.

A closer examination of the reading enjoyment mean scores shows a slight overall decrease in reading enjoyment for participants in both conditions. In the computer-based ABRACADABRA condition, participants' reading enjoyment dropped between pre-test ($M = 14.06$, $SD = 13.64$) and post-test ($M = 9.47$, $SD = 9.80$), $d = -0.36$, showing a small effect size for decline. In the paper ABRACADABRA condition, participants' reading enjoyment similarly

dropped between pre-test (M = 12.82, SD = 11.98) and post-test (M = 11.24, SD = 9.39), d = -0.12, showing a very small effect size for decline.

Reading enjoyment scores were further analyzed by school in order to investigate whether participants' lower scores at post-test related to the CART version (accurate vs. less accurate) that they completed. In the first school, all participants completed the less accurate CART at pre-test and then completed the accurate one at post-test. In the second school, all participants completed the accurate CART both at pre- and post-test. Since participants in the first school completed both the accurate and less accurate versions of CART, reading enjoyment scores could not be reliably compared. Therefore reading enjoyment scores were only examined in the second school to evaluate the effect size of the accurate version of CART. In the second school, participants' reading enjoyment showed a small decline between pre-test (M = 4.20, SD = 5.43) and post-test (M = 2.50, SD = 2.84), d = -0.13, showing a very small effect size for decline.

Given that there is a non-significant p value for the Test x Condition interaction and insignificant effect sizes for participants' reading enjoyment in both the computer-based and paper ABRACADABRA groups, this suggests that 1) technology does not negatively impact participants' reading enjoyment over paper methods over time and 2) the general decline in participants' reading enjoyment is not statistically significant. The insignificant effect sizes for participants' reading enjoyment in the second school further suggests that the small declines in reading enjoyment with the accurate version of CART are not statistically significant.

In summary, this study presents a well-matched sample that did not differ on a host of extraneous variables. These variables pertained to individual characteristics (e.g. gender and age), maternal characteristics (e.g. mother's education and mother's native language), paternal characteristics (e.g. father's native language), and languages spoken in the family (e.g. mother-

child language and father-child language). This sample also did not differ on a series of pre-test measures assessing participants' reading and spelling skills, reading motivation, reading self-concept and reading enjoyment. At post-test, participants showed improvement on multiple measures including reading and spelling skills, as well as perceptions of difficulty associated with reading. The data largely suggests equivalent growth on these measures across both the computer-based and paper ABRACADABRA conditions.

Brief overall summary of findings – say that a well-matched sample did not differ on a host of extraneous and pre-test measures. They improved on multiple, measures between pre- and post-test but most data suggest equivalent growth across the two conditions of interest.

Discussion

This pre- and post-test randomized control trial study aimed to investigate the following research questions: 1) What are the effects of the computer-based vs. paper ABRACADABRA instruction on students' reading skills (listening comprehension, letter reading and word reading) pre- and post-intervention, 2) What are the effects of the computer-based vs. paper ABRACADABRA instruction on students' spelling skills (letter writing and word spelling) pre- and post-intervention, and 3) What are the changes in students' reading motivation, reading self-esteem and reading enjoyment towards computer-based vs. paper ABRACADABRA reading activities pre- and post-intervention.

Major findings from this study included participants demonstrating growth in reading and spelling skills, and showing reduced perceptions of difficulty associated with reading. There was significant change between pre- and post-test often, but no effect of Condition suggesting change was equivalent across the computer-based and paper ABRACADABRA conditions. This pattern was also borne out in effect size analyses. Given this pattern we cannot attribute any of the

changes directly to intervention, as they may have occurred as a result of maturation effects of home, school, or a combination of these factors.

Following this study's findings, there were no reasons for rejecting the following hypotheses: 1) $H1_0$: Students' reading skills will not be differentially impacted (with no negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention, 2) $H2_0$: Students' spelling skills will not be differentially impacted (with no negative effect of computer over paper) by the computer-based vs. paper ABRACADABRA post-intervention, and 3) $H3_0$: Levels of reading motivation, reading self-esteem and reading enjoyment will not vary (increase/decrease) according to students' literacy attainment post-intervention.

The overall findings demonstrate that there are seemingly equivalent effects of the computer-based and paper ABRACADABRA in facilitating growth in Grade 1 and 2 students' listening comprehension, reading and spelling skills. This highlights that both instructional mediums produced no different effects in raising students' performance in these skill domains.

The non-significant differences between the computer-based and paper delivery of ABRACADABRA aligns with McNally et al.'s (2016) study which also found comparable positive effects of the computer-based (r = .183) and paper ABRACADABRA (r = .231) instruction on students' literacy skills. The difference between this study's findings and that of McNally et al.'s (2016) is that neither the computer-based nor the paper delivery were advantaged over the other in this study, whereas in the 2016 study, the paper delivery showed a considerably larger size of effect and therefore was somewhat more effective in raising students' literacy outcomes.

The significance of the present finding that as both computer-based and paper ABRACADABRA delivery had largely indistinguishable effects on students' reading and spelling skills, is that this provides no evidence for a strong view that technology-based instruction has measurable negative effects on students' literacy skills. Solity (cited in Bloom, 2016) has argued that technology may negatively impact writing skills in particular potentially due to lower levels of cognitive processing and engagement required by technology-based literacy instruction.

Furthermore, these findings may indicate that the medium of instruction in the ABRACADABRA program is secondary in importance compared to its actual curriculum content. The content within both modes of ABRACADABRA (computer-based and paper) are equivalent. As there were no significant differences between both modes of delivery, this is consistent with the pattern of the well-designed and well-balanced ABRACADABRA curriculum effectively promoting lower elementary students' literacy skills in Canada (Savage et al., 2013). However, as noted above, these findings cannot be attributed to ABRACADABRA specifically. It should be noted that children made greater than expected average progress (e.g. mean standard scores on reading and spelling at approximately 100 at pre-test and approximately 110 at post-test). This greater than expected growth might potentially be due to the value added by exposure to evidence-based ABRACADABRA content.

In terms of the social-emotional components of reading, there were comparable small negative effects of the computer-based and paper ABRACADABRA on Grade 1 and 2 students' reading self-concept (reading motivation and reading self-esteem) and reading enjoyment. This may suggest that students did not favour the computer-based ABRACADABRA delivery over the paper one. This finding, like the above findings, should be interpreted with caution. These

changes in reading motivation, reading self-esteem and reading enjoyment cannot be related to ABRACADABRA specifically.

There was a small general decline in students' reading self-concept and reading enjoyment, as shown in non-significant small negative effect sizes for the overall measures. This trend, however, was not observed consistently across all the sub-components of reading self-concept (perceptions of difficulty, competence and attitude) across both conditions. Participants in both conditions displayed higher reading self-competency and found reading to be less difficult at post-test compared to pre-test.

Therefore the general declines are not likely to result solely from the computer-based or paper ABRACADABRA instruction, especially since participants' global reading self-concept was better preserved in the computer-based (0.05 difference) ABRACADABRA condition over the paper (0.11 difference) one at post-test. In other words, students' participation in the computer-based or paper ABRACADABRA program was not a consistent cause for students' slightly lower levels of reading motivation, reading self-esteem and reading enjoyment at post-test.

Reading motivation, reading self-esteem and reading attitude are meaningfully related to reading fluency (Mihandoost, 2012). Students who are motivated readers tend to be better at reading (Hussien, 1999). Students with a positive academic self-concept tend to hold more optimistic attitudes towards reading and feel more competent as readers (Chapman, Tunmer, & Prochnow, 2000). Following on from this, students with more positive attitudes towards reading tend to be more competent readers (Mihandoost, 2012). Altogether, these factors are significant in predicting variance in elementary school students' reading fluency (Mihandoost, 2012).

In this study, there was growth in participants' reading and spelling skills at post-test, but a slight general decline in their reading motivation, reading self-esteem and reading enjoyment. This positive growth in students' literacy skills accompanied by a minor dip in their reading motivation, reading self-esteem and reading enjoyment stands as a finding contrasting that of Hussien (1999), Chapman, Tunmer and Prochnow (2000) and Mihandoost's (2012) study findings. This study's effect sizes are small and non-significant, which may reflect measurement error and so should not be overly interpreted. They should, however, be interpreted in light of quite substantial gains in literacy attainment between pre- and post-test on standardized measures.

This leads to a closer examination of the relationship between participants' global reading self-concept and literacy attainment between pre-test and post-test. Reading self-efficacy is an important factor that is related to children's reading achievement (Wigfield et al., 2012), and so it is necessary to investigate whether these two variables were well connected in this study.

At pre-test, participants' reading self-concept and reading skills had a very small positive correlation (r = .050), whereas their reading self-concept and spelling skills had a very small negative correlation (r = -.037). At post-test, participants' reading self-concept and reading skills had a very small negative correlation (r = -.062), as did their reading self-concept and spelling skills (r = -.012).

Contrasting these relations, it appears that at pre-test participants' reading self-concept and reading skills were well connected but their reading self-concept and spelling skills were not. This may suggest that participants' self-perceptions of reading ability are not well aligned with, or accurately reflect, their literacy attainment. These inconsistencies between self-concept ratings and literacy attainment may be due to young children not yet being able to utilize temporal

comparisons of self (Harter, 2012). This suggests that participants may not have the ability to fully recognize growth in reading skills and compare this growth to their reading skills at pre-test. Another possible reason for these inconsistencies could be the experience of learning and working hard as a result of their ABRACADABRA participation. It is rather difficult for young children to recognize that they can simultaneously have attributes or qualities of opposing natures – for instance, good and bad or kind and mean (Harter, 2012). In the context of this study, this could mean that children's perceptions of a challenging situation is related to lower abilities to effectively handle that situation. Participants may therefore have come to think of learning to read as a challenging experience (e.g. involves hard work and long periods of focus) and therefore accordingly lowered their global self-perceptions of reading. Careful interpretation of the correlative relationships between reading self-concept and literacy attainment is needed, as the strength of these relations (ranging between less than 0.1 to 0.3) are very small.

Further interpretation of the relationship between participants' reading self-concept and their literacy gains requires careful consideration of the key components of reading self-concept. These components are perceptions of difficulty associated with reading, competency in completing reading tasks, and attitudes towards reading. As stated earlier in the results, this study found that as students' reading improved, perceptions of reading difficulty were significantly reduced in both the computer-based and paper ABRACADABRA conditions. This finding is consistent with that of Chen and Savage's (2014) study where students' improved reading led to reduced perceptions of reading difficulty. This may suggest that as reading improves, the perceptions of difficulty aspect of reading self-concept improves. Caution should be taken in the interpretation of this result given the absence of controls, however as reading improves reading motivation generally improves. In addition, separate correlational analyses of the other two

reading self-concept components (reading competency and reading attitudes) and literacy attainment were conducted, but correlational strengths were too small to be meaningfully interpreted.

Combining the further interpretations of the relationships between global reading self-concept and literacy attainment, as well as between reading difficulty and literacy attainment, provides a more comprehensive picture of participants' reading self-concept and literacy attainment. It appears that, overall, participants found reading to be less difficult and seemed to show more positive reading self-concept and reading attitudes at post-test. Their global reading self-concept, however, was generally lowered and became more distantly associated with their enhanced post-test literacy attainment on measures of reading and spelling. This seems to point to participants' young age and relative lack of accurate awareness of improvement in reading abilities after their ABRACADABRA participation, which appear to align with difficulties children may have with providing precise self-perceptions aligned with reality (Harter, 2012). Overall, these interpretations should be cautiously maintained given the small correlational strengths of these relationships in this study.

In terms of reading enjoyment, this finding requires interpretation with respect to the CART version used with participants. There were two versions of the CART. In the first version, the researcher only included real authors and no foil authors. In the second corrected version, the researcher included both real authors and foil authors. Participants in the first school were administered the first version of CART at pre-test and then the second corrected version at post-test. This means it was more likely for participants to have scored higher on the first version of CART (with only real authors), and to have scored lower on the second corrected version of this measure (with a balance of real and foil authors).

Analysis of CART scores for participants in the second school, where the second corrected version was consistently administered at pre- and post-test revealed only small to zero effect size for decline. This suggests that with the corrected version of CART, participants' reading enjoyment only showed very small, non-significant declines. Therefore, this points to CART measurement effects (between the first and second versions), possibly at play and being responsible for some portion of decline in participants' reading enjoyment. As with some of the other analyses, due to small non-significant effect sizes, there may be other contributing factors (e.g. school and home) towards lower reading enjoyment, and therefore results should be interpreted with caution.

This study's strengths lie in its design with regards to the high quality, well-delivered ABRACADABRA interventions. Firstly, the pace of lessons was effectively accounted for in that they were consistently timed, monitored and recorded for all ABRACADABRA sessions over the eight-week intervention period. Overall pace of lessons was kept constant across both ABRACADABRA conditions for both schools. Secondly, treatment integrity evaluations revealed that overall intervention delivery in both schools reached normative inter-rater reliability standards typically used in reading intervention studies. Moreover, treatment integrity levels were positively comparable between the computer-based and paper ABRACADABRA conditions, with very minor differences based on the additional evaluation of one computer-based group over the paper groups.

Limitations

This study has a few limitations that will be addressed and then discussed in relation to possible future research directions. Firstly, this study only had two groups (computer-based ABRACADABRA and paper ABRACADABRA), lacking a third control group consisting of

regular classroom instruction. With a third regular classroom instruction group, this would serve as a control and provide additional data on the effectiveness of ABRACADABRA's two delivery modes (computer and paper) in contrast to business as usual classroom instruction. Secondly, this study had a modest sample size of 34 participants. As it stands, the obtained effect sizes were consistent with the corresponding statistical significance values in showing only very small differences by condition. With a larger sample size of participants, effect sizes resulting from participation in the different modes of ABRACADABRA would be more sensitive to the extent of ABRACADABRA's effectiveness in facilitating literacy growth, though formal prospective power calculations showed that the study here was sufficiently powered to detect small-to-medium effects of intervention. Thirdly, this study was run for a total of eight weeks, which is a typical length of time for many published reading interventions. To further enhance the effectiveness of ABRACADABRA on students' literacy outcomes, extending the intervention period to twenty weeks, as in McNally et al.'s (2016) study, could potentially be advantageous. Lastly, the inconsistent administration of the CART for the first school of participants may have produced some measurement errors for participants' reading enjoyment. To ensure accurate investigation of reading enjoyment, providing all participants with the same version of CART both at pre- and post-test is key in maximizing measurement reliability and validity. This, in turn, will enhance this measure's internal validity in investigating ABRACADABRA effects on reading enjoyment.

Conclusions and Future Directions

In conclusion, the biggest contribution of this study is it shows that using a well-designed RCT assessing impacts across a range of academic and related socio-emotional and behaviour measures, there was no evidence of negative effects of computer-based versus paper

ABRACADABRA modality. The main take-away message from this study is that technology-based instruction does not necessarily have negative carryover effects on students' literacy skills, in particular their writing abilities, their motivation or self-concept.

Future research directions could explore the recommendations following from this study's limitations. In a future study investigating the effectiveness of ABRACADABRA's two delivery modes (computer vs. paper), implementing the ABRACADABRA program over a longer duration, including a larger sample size of lower elementary school students, including a delayed post-test, and incorporating a third control group of regular classroom instruction would make for a more methodologically sound study. This prospective study would provide more robust results on the ways in which the two modes of ABRACADABRA (computer vs. paper) differ in their effectiveness compared to regular classroom instruction in terms of facilitating positive literacy growth in students.

References

Abrami, P. C., Borohkovski, E., & Lysenko, L. (2015). The effects of ABRACADABRA on reading outcomes: A meta-analysis of applied field research. *Journal of Interactive Learning Research*, *26*(4), 337-367.

Abrami, P. C., Wade, C. A., Lysenko, L., Marsh, J., & Gioko, A. (2016). Using educational technology to develop early literacy skills in Sub-Saharan Africa. *Education and Information Technologies*, *21*(4), 945-964.

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology*, *84*(4), 489.

Archer, K., Savage, R., Sanghera-Sidhu, S., Wood, E., Gottardo, A., & Chen, V. (2014). Examining the effectiveness of technology use in classrooms: A tertiary meta-analysis. *Computers & Education*, *78*, 140-149.

Aydemir, Z., & Oeztuerk, E. (2012). The effects of reading from the screen on the reading motivation levels of elementary 5th graders. *TOJET: The Turkish Online Journal of Educational Technology*, *11*(3).

Baker, L., Dreher, M. J., & Guthrie, J. T. (2000). Why teachers should promote reading engagement. In L. Baker, M. J. Dreher, & J. T. Guthrie (Eds.), Engaging young readers: Promoting achievement and motivation (pp. 1-16). New York: The Guilford Press.

Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, *34*(4), 452-477.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral

change. *Psychological Review*, *84*(2), 191.

Barnyak, N. C., & McNelly, T. A. (2016). The literacy skills and motivation to read of children enrolled in Title I: A comparison of electronic and print nonfiction books. *Early Childhood Education Journal*, *44*(5), 527-536.

Bavelier, D., Green, C. S., & Dye, M. W. (2010). Children, wired: For better and for worse. *Neuron*, *67*(5), 692-701.

Becker, H. J. (1992). Mindless or mindful use of integrated learning systems. *International Journal of Educational Research*, *21*(1), 65–79.

Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology*, *102*(4), 773.

Berninger, V. W., Abbott, R. D., Augsburger, A., & Garcia, N. (2009). Comparison of pen and keyboard transcription modes in children with and without learning disabilities. *Learning Disability Quarterly*, *32*(3), 123-141.

Blok, H., Oostdam, R., Otter, M. E., & Overmatt, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research*, *72*(1), 101–130.

Bloom, A. (2016, October 28). The pen is mightier than the computer for learning. The Times.

Capin, P., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). Examining how treatment fidelity is supported, measured, and reported in K–3 reading intervention research. *Educational Psychology Review*, 1-35.

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-

concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, *87*(1), 154.

Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, *67*(3), 279-291.

Chapman, J. W., Tunmer, W. E., & Prochnow, E. (2000). Study about reading skills, self-concept and progress in school self-concept. *Journal of Educational Psychology, 92,* 703-708.

Chaudhry, A. S. (2014). Student response to e-books: study of attitude toward reading among elementary school children in Kuwait. *The Electronic Library*, *32*(4), 458-472.

Chen, V., & Savage, R. S. (2014). Evidence for a simplicity principle: teaching common complex grapheme-to-phonemes improves reading and motivation in at-risk readers. *Journal of Research in Reading*, *37*(2), 196-214.

Cheung, A., Mak, B., Abrami, P. C., Wade, A., & Lysenko, L. (2016). The effectiveness of the ABRACADABRA (ABRA) web-based literacy program on primary school students in Hong Kong. *Journal of Interactive Learning Research*, *27*(3), 219-245.

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*(3), 198-215.

Cheung, A. C., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, *48*(3), 277-299.

Ciampa, K. (2014). *Mobile books: Effect of engagement on students' motivation and cognitive*

strategy use (Unpublished doctoral dissertation). Brock University, Canada.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*, 445-459.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Fives, A., Kearns, N., Devaney, C., Lyons, R., Kearns, N., Eaton, P., O'Brien, A., Canavan, J., ... & Russell, D. W. (2013). Evaluation of the effectiveness of Barnardos Wizards of Words reading programme.

Grant, A., Wood, E., Gottardo, A., Evans, M. A., Phillips, L., & Savage, R. S. (2012). Assessing the content and quality of commercially available reading software programs: Do they have the fundamental structures to promote the development of early reading skills in children? *NHSA Dialog*, *15*(4), 319-342.

Greenlee-Moore, M. E., & Smith, L. L. (1996). Interactive computer software: The effects on young children's reading achievement. *Reading Psychology: An International Quarterly*, *17*(1), 43-64.

Grimshaw, S., Dungworth, N., McKnight, C., & Morris, A. (2007). Electronic books: Children's reading and comprehension. *British Journal of Educational Technology*, *38*(4), 583-599.

Gross, S. (2004). Pathways to lasting self-esteem. Bloomington, IN: House.

Guthrie, J. T., Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., ... &

Mitchell, A. M. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, *31*(3), 306-332.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, Volume III* (pp. 403-422). London: Lawrence Erlbaum Associates.

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, *3*(3), 231-257.

Harter, S. (2012). *The construction of the self: Developmental and sociocultural foundations*. New York, NT: Guilford Press.

Hellmich, F., & Hoya, F. (2017). Primary school students' implicit theories and their reading motivation. *Zeitschrift für Psychologie, 225*(2), 117-126.

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Eds.), Intrinsic and extrinsic motivation: The search for optimal motivation and performance (pp. 309–339). San Diego, CA: Academic Press.

Hisken, L. J. (2012). *The correlation between self-esteem and student reading ability, reading level, and academic achievement* (Unpublished master's thesis). University of Central Missouri, United States of America.

Hussien, M. G. (1999). The relationship between motivation to read and reading achievement in grades kindergarten through third (Unpublished master's thesis). Walden University, United States of America.

Jones, T., & Brown, C. (2011). Reading engagement: A comparison between e-books and traditional print books in an elementary classroom. *International Journal of Instruction*, *4*(2), 5-22.

Kiefer, M., Schuler, S., Mayer, C., Trumpp, N. M., Hille, K., & Sachse, S. (2015). Handwriting or typewriting? The influence of pen-or keyboard-based writing training on reading and writing performance in preschool children. *Advances in Cognitive Psychology*, *11*(4), 136.

King, K. A., Vidourek, R. A., Davis, B., & McClellan, W. (2002). Increasing self-esteem and school connectedness through a multidimensional mentoring program. *Journal of School Health*, *72*(7), 294-299.

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, *42*(2), 7-19.

Kucirkova, N., Littleton, K., & Cremin, T. (2017). Young children's reading for pleasure with digital books: Six key facets of engagement. *Cambridge Journal of Education*, *47*(1), 67-84.

Kulik, J. A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say*. Arlington, VA: SRI International.

Kulik, C. L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An update analysis. *Computers in Human Behavior*, *7*(1–2), 75–94.

Liaupsin, C. J., Ferro, J. B., & Umbreit, J. (2012). Treatment integrity in intervention research: Models, measures, and future directions. In *Classroom Behavior, Contexts, and Interventions* (pp. 301-322). Emerald Group Publishing Limited.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A

multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646-656.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133-163.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59-77.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality*, *74*(2), 403-456.

McNally, S., Ruiz-Valenzuela, J., & Rolfe, H. (2016, October). ABRA: Online reading support (Rep.).doi:https:educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EEF_Project_report_ABRA.pdf.

Mihandoost, Z. (2012). Relationship between motivation, self-concept, attitude and fluency of elementary school students. *Life Science Journal*, *9*(4), 9.

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, *48*(6), 1315-1346.

Moody, A. K., Justice, L. M., & Cabell, S. Q. (2010). Electronic versus traditional storybooks: Relative influence on preschool children's engagement and communication. *Journal of Early Childhood Literacy*, *10*(3), 294-313.

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.

Newlin, R. B. (2003). Paws for reading: An innovative program uses dogs to help kids read better. *School Library Journal*, *49*(6), 43.

Newman, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F.M. Newman (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11-39). New York: Teachers College.

Norton, T. L., & Anfin, C. S. (2003). Birthday booktalks: Fostering emergent literacy and self-esteem in young children. *Reading Horizons*, *43*(3), 203.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33-84.

Ouellette, G., & Tims, T. (2014). The write way to spell: printing vs. typing effects on orthographic learning. *Frontiers in Psychology*, *5*, 117.

Ouyang, J. (1993). *Meta-analysis: CAI at the level of elementary education*. Paper presented at the World Conference on Education Multimedia and Hypermedia.

Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept–Testing the reciprocal effects model. *Learning and Instruction*, *29*, 21-30.

Savage, R. S., Abrami, P. C., Piquette, N. A., Wood, E., Deleveaux, G., Sanghera-Sidhu, S., &

Burgos, G. (2013). A (Pan-Canadian) cluster randomized control effectiveness trial of the ABRACADABRA web-based literacy program. *Journal of Educational Psychology*, *105*(2), 310.

Savage, R. S., & Cloutier, E. (2017). Early reading interventions. *Theories of Reading Development*, *15*, 409.

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

Sénéchal, M., LeFevre, J. A., Hudson, E., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology*, *88*(3), 520.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407-441.

Selwyn, N. (2012). *Education in a digital world: Global perspectives on technology and education*. Routledge.

Soe, K., Koki, S., & Chang, J. M. (2000). *Effect of computer assisted instruction (CAI) on reading achievement: A meta-analysis*. Washington, DC: OERI.

Tracey, D. H., & Morrow, L. M. (2006). *Lenses on reading: An introduction to theories and models*. New York: The Guilford Press.

Valm, A. (2017, June 14). 100 best Canadian kids' books of all time. Retrieved November 15, 2017, from https://www.todaysparent.com/family/books/top-100-canadian-kids-books/.

Vandewater, E. A., Bickham, D. S., & Lee, J. H. (2006). Time well spent? Relating television use to children's free-time activities. *Pediatrics*, *117*(2), e181-e191.

Wang, J. H. Y., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic

motivation, amount of reading, and past reading achievement on text comprehension between US and Chinese students. *Reading Research Quarterly*, *39*(2), 162-186.

Weis, R., & Cerankosky, B. C. (2010). Effects of video-game ownership on young boys academic and behavioural functioning: A randomized controlled study. *Psychological Science, 21*, 463-470. Doi:10.1177?0956797610362670

Wigfield, A., Cambria, J., & Ho, A. N. (2012). Motivation for reading information texts. *Adolescents' Engagement in Academic Literacy*, 52.

Wilkinson, G. S., & Robertson, G. J. (2006). *WRAT 4: Wide Range Achievement Test*. Lutz, FL: Psychological Assessment Resources.

Williams, K. T. (2001). *Group reading assessment and diagnostic evaluation*. American Guidance Service.

Wolgemuth, J. R., Abrami, P. C., Helmer, J., Savage, R. S., Harper, H., & Lea, T. (2014). Examining the impact of ABRACADABRA on early literacy in Northern Australia: An implementation fidelity analysis. *The Journal of Educational Research*, *107*(4), 299-311.

Wolgemuth, J. R., Savage, R. S., Helmer, J., Harper, H., Lea, T., Abrami, P. C., ... & Louden, W. (2013). ABRACADABRA aids Indigenous and non-Indigenous early literacy in Australia: Evidence from a multisite randomized controlled trial. *Computers & Education*, *67*, 250-264.

Wollscheid, S., Sjaastad, J., & Tømte, C. (2016). The impact of digital devices vs. Pen (cil) and paper on primary school students' writing skills–A research review. *Computers & Education*, *95*, 19-35.

Wollscheid, S., Sjaastad, J., Tømte, C., & Løver, N. (2016). The effect of pen and paper or tablet computer on early writing–a pilot study. *Computers & Education*, *98*, 70-80.

Wood, E., Grant, A. K., Gottardo, A., Savage, R. S., & Evans, M. A. (2017). Software to promote young children's growth in literacy: A comparison of online and offline formats. *Early Childhood Education Journal*, *45*(2), 207-217.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, *49*(2), 156.

Appendix A

Computer-based ABRACADABRA Curriculum (Weeks 1-6)

WEEK 1: Text – I Can Move Like A….

DAY 1		
Intro	**Navigation**	**3 minutes**
Intro	**Group Rules (S*T*A*R)**	**5 minutes**
Intro	**Intro to ABRA (characters, stories)**	**7 minutes**
DAY 2		
Intro	**Recap Group Rules**	**2 minutes**
Intro	**Navigation Strategy**	**4 minutes**
D	**Letter Bingo**	**9 minutes**
DAY 3		
D	**Animated Alphabet – S and P**	**3 minutes**
D	**Auditory Blending – Level 1 (2 phonemes)**	**2 minutes**
R	**Tracking – First half of book**	**8 minutes**
D	**Auditory Blending – Level 1 (2 phonemes)**	**2 minutes**
DAY 4		
D	**Animated Alphabet – S and P**	**1 minute**
D	**Auditory Blending – Level 1**	**2 minutes**
R	**Tracking – Second half of the book**	**8 minutes**
D	**Auditory Blending – Level 1**	**2 minutes**
D	**Matching Sounds – Level 1**	**2 minutes**

Remember, when the rate of correct response is 80-90% (as a group) for three consecutive entries into a particular activity you can move on to the next level.

Important Please Note:
Here and throughout the intervention, it may be that some children move very quickly from one level to the nest so the guidelines are approximate only, we should never hold children back. 'Baseline' performance is assumed here, hence assessment involves starting with letter names to be sure they are secure and then moving to letter sounds. However, if children do not know letter names then of course we would stay here. Letter sounds are in frequency order. Many children will know some so start at the point where they do not know any. When children know all sounds taught in ABRA (including vowel digraphs) they can use the time for Reading and Comprehension activities. Equally, when children have reached the ceiling on Basic Decoding/Blending Train/and Word Changing they should spend time on Tracking and related activities for reading fluency.

WEEK 2: Text – I Can Move Like A….

DAY 5		
D	Animated Alphabet – S,P,M	3 minutes
D	Auditory Blending (Level 2 if 80-90% accuracy achieved at Level 1)	2 minutes
R	Tracking – First half of book	8 minutes
D	Auditory Blending (Level 2 if 80-90% accuracy achieved at Level 1)	2 minutes
DAY 6		
D	Animated Alphabet – S and T	3 minutes
D	Blending Train – Level 1 or 2	2 minutes
R	Tracking – Second half of book	8 minutes
D	Blending Train – Level 1 or 2	2 minutes
DAY 7		
D	Animated Alphabet – P and T	3 minutes
D	Auditory Blending – Level 1 or 2	2 minutes
C	Vocabulary	8 minutes
D	Auditory Blending – Level 1 or 2	2 minutes
DAY 8		
C	Vocabulary (Finish)	6 minutes
D	Blending Train (Level 2 if 80-90% accuracy achieved at Level 1)	2 minutes
R	Tracking (Computer reads, children join in whole story)	5 minutes
D	Blending Train (Level 2 if 80-90% accuracy achieved at Level 1)	2 minutes

WEEK 3: Text – How A Bean Sprouts

DAY 9		
D	Animated Alphabet - S,P,M,T	3 minutes
D	Blending Train – Level 1/2	3 minutes
R	Tracking (Computer reads child choral co-reads, first 3-4 pages)	6 minutes
D	Blending Train – Level 1/2	3 minutes
DAY 10		
D	Animated Alphabet - L,R,M	3 minutes
D	Auditory Blending – Level 2	3 minutes
R	Tracking (Computer read child shares as above, last 3-4 pages)	6 minutes
D	Auditory Blending – Level 2	3 minutes
DAY 11		
D	Animated Alphabet – L,R,M	3 minutes
D	Auditory Blending – Level 2/3	3 minutes
C	Vocabulary	7 minutes
D	Auditory Blending – Level 3	2 minutes
DAY 12		
D	Animated Alphabet – L,R,M,I	3 minute
D	Auditory Blending – Level 3	2 minutes
C	Vocabulary	8 minutes
D	Auditory Blending – Level 3	2 minutes

WEEK 4: Text – How A Bean Sprouts

DAY 13		
D	**Blending Train – Level 2/3**	**2 minutes**
D	**Animated Alphabet – L,R,M,I**	**3 minutes**
C	**Comprehension Monitoring (first 3-4 pages)**	**8 minutes**
D	**Blending Train – Level 2/3**	**2 minutes**
DAY 14		
D	**Animated Alphabet - L,I,N**	**2 minutes**
D	**Auditory Blending – Level 3/4**	**2 minutes**
C	**Comprehension Monitoring (last 3-4 pages)**	**9 minutes**
D	**Auditory Blending – Level 3/4**	**2 minutes**
DAY 15		
D	**Animated Alphabet – I,D,N**	**2 minutes**
D	**Blending Train – Level 3**	**2 minutes**
C	**Sequencing – Level 1 (3 cards)**	**9 minutes**
D	**Blending Train – Level 3**	**2 minutes**
DAY 16		
D	**Animated Alphabet – D,I,R**	**2 minutes**
D	**Auditory Blending – Level 4**	**2 minutes**
C	**Sequencing – Level 2 (5 cards)**	**6 minutes**
D	**Blending Train – Level 4**	**2 minutes**
	Student Choice	**3 minutes**

Please remember that any time there is a Student Choice activity it is limited to activities that students have already experienced.

WEEK 5: Text – How A Bean Sprouts

DAY 17		
D	Animated Alphabet – K,I,B	2 minutes
D	Blending Train – Level 3/ 4 (continue if not reached 90% accuracy)	2 minutes
C	Vocabulary (review all)	6 minutes
D	Blending Train – Level 3/4 (same criteria as above)	4 minutes
D	Basic Decoding – Level 1/2	2 minutes
DAY 18		
D	Animated Alphabet – L,I,B	3 minutes
D	Auditory Blending – Level 4/5	2 minutes
C	Comprehension Monitoring (all text)	8 minutes
D	Basic Decoding – Level 1/2	2 minutes
DAY 19		
D	Animated Alphabet – R,I,B	2 minutes
D	Auditory Blending – Level 4/5	2 minutes
C	Sequencing – Level 1 (3 page version)/ Level 2 (5 page version)	7 minutes
D	Basic Decoding – Level 2/3	4 minutes
DAY 20		
D	Letter-Sound Search – Level 1	5 minutes
D	Basic Decoding – Level 2/3	2 minutes
R	Tracking (all text, child choral co-reads)	8 minutes

WEEK 6: Text – The Dove and the Ant

DAY 21		
D	Animated Alphabet – SH,A	2 minutes
D	Blending Train – Level 4/5	3 minutes
C	Prediction (first 3rd of story)	7 minutes
D	Blending Train – Level 4/5	3 minutes
DAY 22		
D	Animated Alphabet – I,SH,A	3 minutes
D	Basic Decoding – Level 3	3 minutes
C	Prediction (second 3rd of story)	6 minutes
D	Basic Decoding – Level 3	3 minutes
DAY 23		
D	Animated Alphabet – SH,A,L	2 minutes
D	Auditory Blending – Level 5/6	2 minutes
C	Prediction (last 3rd of story)	7 minutes
D	Auditory Blending – Level 5/6	2 minutes
R	High Frequency Words – Level 1	2 minutes
DAY 24		
D	Animated Alphabet – SH,A,I,M	2 minutes
D	Basic Decoding – Level 3/4	2 minutes
C	Story Response	7 minutes
D	Basic Decoding – Level 3/4	2 minutes
	Student Choice	2 minutes

Appendix B

Paper ABRACADABRA Curriculum (Weeks 1-6)

OVERVIEW WEEK 1

Book: I Can Move Like A…….

DAY 1		
Intro	Introduction to the Group	3 minutes
Intro	Group Rules	6 minutes
Intro	Letters in Name Activity	6 minutes
DAY 2		
Intro	Recap Group Rules	2 minutes
Intro	Letters in Name Activity	4 minutes
D	Letter Bingo	9 minutes
DAY 3		
D	Alphabet	3 minutes
D	Blending	2 minutes
R	Tracking	8 minutes
D	Blending	2 minutes
DAY 4		
D	Alphabet	1 minute
D	Blending	2 minutes
R	Tracking	8 minutes
D	Blending	2 minutes
D	Same Sounds	2 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

OVERVIEW WEEK 2

Book: I Can Move Like A…….

DAY 5		
D	Alphabet	3 minutes
D	Blending	2 minutes
R	Tracking	8 minutes
D	Blending	2 minutes
DAY 6		
D	Alphabet	3 minutes
D	Blending	2 minutes
R	Tracking	8 minutes
D	Blending	2 minutes
DAY 7		
D	Alphabet	3 minutes
D	Blending	2 minutes
C	Vocabulary	8 minutes
D	Blending	2 minutes
DAY 8		
C	Vocabulary	6 minutes
D	Blending	2 minutes
R	Tracking	5 minutes
D	Blending	2 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

OVERVIEW WEEK 3

Book: How a Bean Sprouts

DAY 9		
D	Alphabet	3 minutes
D	Blending	3 minutes
R	Tracking	6 minutes
D	Blending	3 minutes
DAY 10		
D	Alphabet	3 minutes
D	Blending	3 minutes
R	Tracking	6 minutes
D	Blending	3 minutes
DAY 11		
D	Alphabet	3 minutes
D	Blending	3 minutes
C	Vocabulary	7 minutes
D	Blending	2 minutes
DAY 12		
D	Alphabet	3 minute
D	Blending	2 minutes
C	Vocabulary	8 minutes
D	Blending	2 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

OVERVIEW WEEK 4

Book: How a Bean Sprouts

DAY 13		
D	Blending	2 minutes
D	Alphabet	3 minutes
C	Monitoring	8 minutes
D	Blending	2 minutes
DAY 14		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Monitoring	9 minutes
D	Blending	2 minutes
DAY 15		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Sequencing	9 minutes
D	Blending	2 minutes
DAY 16		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Sequencing	9 minutes
D	Blending	2 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

OVERVIEW WEEK 5

Book: How a Bean Sprouts

DAY 17		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Vocabulary	6 minutes
D	Decoding	5 minutes
DAY 18		
D	Alphabet	3 minutes
D	Blending	2 minutes
C	Monitoring	8 minutes
D	Decoding	2 minutes
DAY 19		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Sequencing	7 minutes
D	Decoding	4 minutes
DAY 20		
D	Letter-Sound Search	5 minutes
D	Decoding	2 minutes
R	Tracking	8 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

OVERVIEW WEEK 6

Book: The Dove and the Ant

DAY 21		
D	Alphabet	2 minutes
D	Blending	3 minutes
C	Prediction	7 minutes
D	Blending	3 minutes
DAY 22		
D	Alphabet	3 minutes
D	Decoding	3 minutes
C	Prediction	6 minutes
D	Decoding	3 minutes
DAY 23		
D	Alphabet	2 minutes
D	Blending	2 minutes
C	Prediction	7 minutes
D	Blending	2 minutes
D	High Frequency Words	2 minutes
DAY 24		
D	Alphabet	2 minutes
D	Decoding	2 minutes
C	Story Response	9 minutes
D	Decoding	2 minutes

D = decoding tasks
C = comprehension tasks
R = reading tasks

Appendix C

Parent or Legal Tutor Information Letter and Consent Form

Parent or Legal Tutor

A Randomized Control Trial of Computer-Based vs. Paper and Pencil ABRACADABRA: Effects on Grade 1 and Grade 2 Ready and Spelling Skills, Reading Motivation, Reading Self-Esteem and Reading Enjoyment

Dear Parent or Legal Tutor:

My name is Kristen Sha and I am a Master's thesis student in Educational Psychology – Human Development within the Educational and Counselling Psychology Department at McGill University. My thesis project is being supervised by Dr. Tara Flanagan. I am contacting you to request your permission for your child to participate in my research study.

The aim of this study is to help us understand the effects of the delivery format of a research, evidence-based reading intervention, on the reading and spelling abilities, reading motivation, reading self-esteem and reading enjoyment of all Grade 1 and Grade 2 students.

At the beginning of this study, your child will complete literacy and general ability assessments to determine his or her initial literacy abilities and general abilities. The total expected time for assessment is approximately 125 minutes at pre-test and approximately 100 minutes at post-test. There will be a total of 3 sittings each for the administration of these assessments at pre-test and post-test. These assessments will be conducted in a quiet area as close to your child's classroom as possible. Students will complete assessments one-on-one, and whenever possible, in small groups, with the researcher. During assessment sessions the classroom teacher will be working with students in the classroom as usual. I will work closely with the classroom teacher to arrange an optimal schedule to assess students, so that on average they can still participate in the regular variety of lessons in a typical school day. Parents will also be asked to complete a brief questionnaire on your child's health and learning background and some parent demographic information.

The main part of this study will involve your child's participation in a reading intervention, which will take place three to four times per week, for approximately 20 to 25 minutes each session during the regular school day. The intervention will take place in small groups facilitated by the researcher in a quiet instructional space close to your child's classroom. I will work closely with the classroom teacher to schedule the intervention sessions so that your child is able to physically remain in the classroom as much as possible. The study will take place for 8 consecutive weeks. At the end of 8 weeks, your child's literacy skills will be tested again to determine his or her literacy skills post-intervention.

Only my research supervisor at McGill, Dr. Tara Flanagan, and my co-supervisor at University College London, Dr. Rob Savage, an undergraduate research assistant and I will have access to your child's information and data. All information that is gathered in this study is

confidential so your child's name will not be attached with the information. At the end of this study, results will be reported as anonymous group data in academic journals and at academic conferences.

I would like to share your child's results to his or her teacher so that it can further inform best instructional classroom practices. It is only with your consent that your child's results will be shared with his or her classroom teacher. This additional consent is completely optional, and your decision will not affect your child's eligibility to participate in this study.

Participation is entirely at your discretion, and you are free to not provide consent. I will explain the study to your child and ask for their consent to participate in this study. I will also explain to your child that they can stop participating in the study at any time. There are no foreseeable risks attached to participating in this study. Agreeing for your child to participate in this study does not waive any of your rights or release the researchers from their responsibilities. To ensure the study is being conducted properly, authorized individuals such as a member of the Research Ethics Board, may have access to your child's research information. By signing this consent form, you are allowing such access. You may choose to terminate your child's participation at any time. The decision to withdraw your child's participation will not affect the relationships you or your child have with the school, the researcher or McGill University. If you do not wish for your child to participate in this study, your child will continue with their day-to-day classroom activities as usual.

Should you have any questions or require further information, please email me:
Kristen Sha, kristen.sha@mail.mcgill.ca, (647) 883-9913

If you have any questions or concerns regarding your child's rights or welfare as a participant in this study, you may contact the Research Ethics Officer, Lynda McNeil, at (514) 398-6831 or by e-mail lynda.mcneil@mcgill.ca.

Sincerely,

Kristen Sha
MA Candidate
Faculty of Education, McGill University

Dr. Tara Flanagan
Professor & Graduate Advisor
Faculty of Education, McGill University

A Randomized Control Trial of Computer-Based vs. Paper and Pencil ABRACADABRA: Effects on Grade 1 and Grade 2 Ready and Spelling Skills, Reading Motivation, Reading Self-Esteem and Reading Enjoyment

Please indicate whether you give permission for your child to participate in this study by ticking off one of the below statements, signing your name, and asking your child to sign his or her name.

Keep the first part of this letter for your own records, and kindly return the second half of the consent form via your child to the researcher.

_____ Yes, I give permission for my child to participate in this study.

_____ Yes, I give permission for my child's results to be shared with his or her classroom teacher. (optional additional consent)

_____	_____
Signature of Parent/Legal Tutor	Printed Name of Parent/Legal Tutor
_____	_____
Printed Name of Child	Date

Appendix D

Principals/Teachers Information Letter and Consent Form

Information Letter for Principals/Teachers

A Randomized Control Trial of Computer-Based vs. Paper and Pencil ABRACADABRA: Effects on Grade 1 and Grade 2 Ready and Spelling Skills, Reading Motivation, Reading Self-Esteem and Reading Enjoyment

Dear Principal and Teachers:

My name is Kristen Sha and I am a Master's thesis student in Educational Psychology – Human Development within the Educational and Counselling Psychology Department at McGill University. My thesis project is being supervised by Dr. Tara Flanagan. I am contacting you to request your permission for your school to participate in my research study.

The aim of this study is to help us understand the effects of the delivery format of a research, evidence-based reading intervention, on the reading and spelling abilities, reading motivation, reading self-esteem and reading enjoyment of all Grade 1 and Grade 2 students. The reading intervention will be delivered in two different formats: computer and paper-and-pencil.

At the beginning of this study, your student will, with prior obtained parental consent, complete literacy and general ability assessments to determine his or her initial literacy abilities and general abilities. The total expected time for assessment is approximately 125 minutes at pre-test and approximately 100 minutes at post-test. There will be a total of 3 sittings each for the administration of these assessments at pre-test and post-test. These assessments will be conducted in a quiet area as close to students' classrooms as possible. Students will complete assessments one-on-one, and whenever possible, in small groups, with the researcher. During assessment sessions the classroom teacher will be working with students in the classroom as usual. I will work closely with the classroom teacher to arrange an optimal schedule to assess students, so that on average they can still participate in the regular variety of lessons in a typical school day.

The main part of this study will involve your student's participation in a reading intervention, which will take place three to four times per week, for approximately 20 to 25 minutes each session during their homeroom period. The intervention will take place in small groups facilitated by the researcher in a quiet instructional space close to students' classrooms. I will work closely with the classroom teacher to schedule the intervention sessions so that students are able to physically remain in the classroom as much as possible. The intervention will be 8 weeks long. At the end of 8 weeks, your student's literacy skills will be tested again to determine his or her literacy skills post-intervention.

I anticipate value and insight to be added to your student's reading development and reading performance, based on encouraging results from previous research studies carried out in schools with lower elementary school age students.

What I am requiring from the school: Your school's participation is key to this research study. For principals, I am asking you to allow this intervention to be conducted in your school for 8 consecutive weeks. I will randomly assign students to one of two ABRACADABRA reading interventions, both of which reflect high quality, evidence-based practices in the field. I wish to work with all children in Grade 1 and Grade 2.

What I am requiring from the teacher: The willingness of teachers to allow the researcher access to his or her students is also vital to the success of this study. I ask that teachers allow the researcher to take small groups of students at one time to a quiet space within the school to run the ABRACADABRA reading interventions during the designated intervention sessions.

I am highly aware and understanding of school structures, practices and needs based on my previous experience working in schools. In light of this, I am committed in my position to be a flexible, considerate and professional collaborator. My priority will be to maintain a collaborative approach at all times and to work around the needs of schools.

This study has been approved by the McGill University Research Ethics Board and by your school board.

Confidentiality: All information gathered for this study will be kept entirely confidential and anonymous. Data will be used for publication in academic journals. Data will be presented anonymously in all instances. Your school's name will not be identified in publications. Data will not be distributed within school boards.

Voluntary Participation: Participation in this study is entirely at your discretion, and you are free to not provide consent. No foreseeable risks are involved in participating in this study.

Withdrawal from the Study: Students at your school can withdraw from the study at any time, or decline to participate for any reason. The decision to end participation on any basis will not impact any relationships with the researcher or McGill University. You may also choose to withdraw your school or class from all or parts of the study.

Questions about the Research: Should you have any questions or concerns, or require further information, please email me: Kristen Sha, kristen.sha@mail.mcgill.ca, (647) 883-9913.

Should you have any ethical concerns regarding this research study, you may contact Lynda McNeil, the Research Ethics Officer of REB-III studies for McGill University, by email at lynda.mcneil@mcgill.ca or by phone at (514) 398-6831.
I hope that your school will be a part of this beneficial research study. I look forward to hearing from you soon.

Sincerely,

Kristen Sha
MA Candidate
Faculty of Education, McGill University

Dr. Tara Flanagan
Professor & Graduate Advisor
Faculty of Education, McGill University

A Randomized Control Trial of Computer-Based vs. Paper and Pencil ABRACADABRA: Effects on Grade 1 and Grade 2 Ready and Spelling Skills, Reading Motivation, Reading Self-Esteem and Reading Enjoyment

Please indicate whether you give permission for your school to participate in this study by ticking off one of the below statements and signing your name.

Keep the first part of this letter for your own records, and kindly return the second half of the consent form to the researcher.

_____Yes, I give permission for my school to participate in this study. (Principals)

_____Yes, I give my permission to participate in this study. (Teachers)

Name of School (please print)

_____

Name of Principal/Teacher (please print)

_____

Signature _____ **Date** _____

Appendix E

Child Assent Script (Computer Version ABRACADABRA)

Child Assent Script (Computer Version ABRACADABRA)

(Grade 1 and Grade 2 Students)

I am doing a study on kids and reading. You will be doing reading games on the computer. We are going to work together and I am going to help you. If you need a break during the study, you can say "I need a break". If you do not want to continue being in the study anymore, you can say "I want to stop" or "I do not want to do this anymore" at any time. Do you have any questions? Is this something you want to do?

(Obtain verbal assent).

Appendix F

Child Assent Script (Paper and Pencil Version ABRACADABRA)

Child Assent Script (Paper & Pencil Version ABRACADABRA)

(Grade 1 and Grade 2 Students)

I am doing a study on kids and reading. You will be doing reading activities. We are going to work together and I am going to help you. If you need a break during the study, you can say "I need a break". If you do not want to continue being in the study anymore, you can say "I want to stop" or "I do not want to do this anymore" at any time. Do you have any questions? Is this something you want to do?

(Obtain verbal assent).

Appendix G

Reading Self-Concept Scale (Sample Items)

Reading Self-Concept Scale (Tunmer & Chapman, 1995)

Sample Items From the Reading Self-Concept Scale (30-Item Version)

Competence Subscale

4. Can you work out hard words by yourself when you read?
9. Are you good at remembering words?
21. Can you work out sounds in words?
27. Do you think you read well?
29. Can you work out hard words in a story even if there are no pictures?

Attitude Subscale

5. Do you like word games in class?
10. Is it fun for you to read books?
19. Do you look forward to reading?
22. Do you like reading to yourself?
30. Do you like reading at home?

Difficulty Subscale

3. Is reading to the class hard for you?
6. Are the books you read in class too hard?
18. Do you make lots of mistakes in reading?
23. Do you need extra help in reading?
26. Is it hard for you to understand the stories you have to read in class?

Received July 21, 1993
Revision received August 8, 1994
Accepted August 26, 1994 ■

Appendix H

Children's Author Recognition Test – Adapted Version

Adapted from Children's Author Recognition Test (CART, Senechal et al., 1996)

Canadian Children's Author Recognition Test

1. Paulette Bourgeois
2. Kenneth Oppel
3. Phoebe Gilman
4. Eugenie Fernandes
5. Walter Lee
6. Julie Flett
7. Ian Wallace
8. Julie Morstad
9. Barbara Reid
10. Becky Benson
11. Jonathan Auxier
12. Melanie Watt
13. Tim Wynne-Jones
14. Celia Barker
15. Kathy Stinson
16. Dennis Lee
17. Robert Munsch
18. Kyo Maclear
19. Jean Little
20. Ruth Ohi
21. Gordon Korman
22. Andres Davis
23. Rick Coleman
24. Jo Hunter
25. Fred Carr
26. Kimberly Miller
27. Kit Pearson
28. Cristina Hardy
29. Farley Mowat
30. L.M. Montgomery
31. Arnold Meyer
32. Marie-Louise Gay
33. Janet Lunn
34. Cary Weaver
35. Lionel Robertson
36. Rudy Bryant
37. Ramiro Watkins
38. Belinda Reese
39. Nichole Webster
40. Jon Klassen